
BACHELORARBEIT

Frau
Anja Grajewski

**Mutationsanalyse für Proteine
basierend auf Energieprofilen**

Mittweida, 2011

BACHELORARBEIT

Mutationsanalyse für Proteine basierend auf Energieprofilen

Autor:

Frau

Anja Grajewski

Studiengang:

Biotechnologie/Bioinformatik

Seminargruppe:

BI08w1-B

Erstprüfer:

Prof. Dr. rer. nat. Dirk Labudde

Zweitprüfer:

Dipl.-Inf. (FH) Daniel Stockmann

Einreichung:

Mittweida, 22.08.2011

Verteidigung/Bewertung:

Mittweida, 25.08.2011

Bibliografische Beschreibung:

Grajewski, Anja:

Mutationsanalyse für Proteine basierend auf Energieprofilen. - 2011. - 8 S., 52 S., 19 S., Mittweida, Hochschule Mittweida, Fakultät Mathematik/ Naturwissenschaften/ Informatik, Bachelorarbeit, 2011

Referat:

Die Bachelorarbeit behandelt die Mutationsanalyse von Proteinen mit Hilfe von Strukturvergleichen und Energieprofilen. Dabei werden die erforderlichen Energieprofile errechnet und die daraus resultierenden Ergebnisse in Hinsicht auf den Wildtyp klassifiziert.

Inhaltsverzeichnis

Inhaltsverzeichnis.....	I
Abbildungsverzeichnis.....	III
Tabellenverzeichnis.....	VII
Abkürzungsverzeichnis.....	VIII
Danksagung.....	IX
1 Einleitung/ Motivation.....	1
1.1 Ziel der Arbeit.....	2
2 Biologische Grundlagen.....	3
2.1 Proteine.....	3
2.1.1 Membranproteine.....	5
2.1.1.2 Bacteriorhodopsin.....	6
2.1.2 Globuläre Proteine.....	11
2.1.2.1 T4 Lysozym	12
2.2 Wechselwirkungskräfte in Proteinen.....	13
2.2.1 Kovalente Bindungen.....	13
2.2.2 Dipol-Dipol-Wechselwirkungskräfte.....	14
2.2.3 Wasserstoffbrückenbindungen.....	14
2.2.4 Van-der-Waals-Kräfte.....	15
2.2.5 Abstoßungskräfte.....	16
2.2.6 Hydrophober Effekt.....	16
2.3 Proteinfaltung.....	17
3 Bioinformatische Grundlagen.....	19
3.1 Biodatenbanken.....	19
3.2 Energieprofile.....	20
3.3 Tools für die Mutationsanalyse.....	24
3.3.1 DaliLite.....	24
3.3.2 PyMOL.....	25
3.3.3 MEPAL.....	25
4 Strukturanalyse.....	29
4.1 Literaturrecherche bezüglich der Mutationsauswirkungen	30
4.2 Strukturvergleich mittels Dali-Server	31

4.3 Strukturvergleich mittels RMSD-Diagramms.....	34
4.4 Strukturvergleich mittels Distanzbaum	37
5 Energieprofilanalyse.....	40
5.1 Analyse mittels Energieprofilalignment	40
5.2 Analyse mittels MEPAL.....	44
5.2.2 Energiedistanzbäume.....	47
6 Diskussion	48
6.1 Vergleich der RMSD- und Mutationsenergieprofil diagramme.....	48
6.2 Vergleich von Distanzbaum und Energieprofil distanzbaum.....	49
6.3 Schlussfolgerung.....	52
Anhang	53
A.1 Mutationswahrscheinlichkeitsanalyse.....	53
A.2 RMSD-Diagramme.....	55
A.2.1 Bacteriorhodopsin	55
A.2.2 T4 Lysozym.....	58
A.3 Energieprofilalignments.....	60
A.3.1 Bacteriorhodopsin.....	60
A.3.2 T4 Lysozym.....	63
A.4 Mutationsenergieprofile.....	66
A.4.1 Bacteriorhodopsin.....	66
A.4.2 T4 Lysozym.....	69
Literaturverzeichnis.....	72
Erklärung zur selbständigen Anfertigung der Arbeit.....	77

Abbildungsverzeichnis

Abbildung 1: Allgemeine Struktur einer Aminosäure.....	3
Abbildung 2: Integrale Membranproteine.....	5
Abbildung 3: Periphere Membranproteine.....	6
Abbildung 4: Karte der Elektronendichte eines zweidimensionalen Bacteriorhodopsin-Gitters.....	7
Abbildung 5: Struktur von Bacteriorhodopsin.....	8
Abbildung 6: Protonenweg innerhalb von Bacteriorhodopsin.....	9
Abbildung 7: Photozyklus von Bacteriorhodopsin.....	10
Abbildung 8: Tertiärstruktur des T4 Lysozym Proteins.....	12
Abbildung 9: Kovalente Bindung zweier Chloratome.....	13
Abbildung 10: Wasserstoffbrückenbindung.....	15
Abbildung 11: Abhängigkeit der Hydrophobizität der Aminosäuren und ihrer Oberfläche.....	17
Abbildung 12: Schematische Darstellung der Proteinfaltung.....	18
Abbildung 13: Multistrukturalignment (Bacetrirhodopsin).....	32
Abbildung 14: Multistrukturalignment (Lysozym).....	33
Abbildung 15: RMSD-Berechnung mit RMSD Calculator. Vergleich von Bacteriorhodopsin und die Mutante 3HAP.....	34
Abbildung 16: Bacteriorhodopsin-RMSD-Diagramm: Wildtyp(2NTU)-Mutante(3HAP).....	35
Abbildung 17: Lysozym-RMSD-Diagramm: Wildtyp(3LZM)-Mutante(3C81)....	36
Abbildung 18: Distanzbaum von Bacteriorhodopsin mit der Ausgabe von Mobyle@Pasteur.....	39
Abbildung 19: Distanzbaum von Lysozym mit der Ausgabe von Mobyle@Pasteur.....	39
Abbildung 20: Energieprofil des Wildtyps Bacteriorhodopsin.....	41
Abbildung 21: Energieprofil der BR-Mutante 3HAP.....	41
Abbildung 22: Energieprofilalignment von 2NTU und 3HAP.....	41
Abbildung 23: Energieprofil des Wildtyps Lysozym.....	42
Abbildung 24: Energieprofil der Mutante 3C81.....	42
Abbildung 25: Energieprofilalignment von 3LZM und 3C81.....	42

Abbildung 26: Mutationsenergieprofil von 2NTU und 3HAP.....	43
Abbildung 27: Mutationsenergieprofil von 3LZM und 3C81.....	43
Abbildung 28: Multienergiealignment von Bacteriorhodopsin.....	45
Abbildung 29: Multienergiealignment von Lysozym.....	46
Abbildung 30: Energieprofilabstandsbau von Bacteriorhodopsin.....	47
Abbildung 31: Energieprofilabstandsbau von Lysozym.....	47
Abbildung 32: Vergleich RMSD-Diagramm und Mutationsprofil von 2NTU-3HAP.....	48
Abbildung 33: Vergleich RMSD-Diagramm und Mutationsprofil von 3LZM-3C81.....	48
Abbildung 34: Vergleich von Abstandsbau und Energieabstandsbau (Bacteriorhodopsin).....	49
Abbildung 35: Wasserstoffbrücken des Wildtyps 2NTU und der Mutante 3HAP	50
Abbildung 36: Vergleich von Abstandsbau und Energieabstandsbau (Lysozym).....	51
Abbildung 37: Wasserstoffbrücken des Wildtyps 3LZM und der Mutante 3C81	51
Abbildung 38: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(2I21).....	55
Abbildung 39: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(2I1X).....	55
Abbildung 40: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(1PXS).....	55
Abbildung 41: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(1PXR).....	56
Abbildung 42: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(1Q5I).....	56
Abbildung 43: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(1P8I).....	56
Abbildung 44: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(2WJL).....	57
Abbildung 45: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(3HAQ).....	57
Abbildung 46: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(1TN0).....	57
Abbildung 47: RMSD-Diagramm: Wildtyp(3LZM)-Mutante(3C7Y).....	58
Abbildung 48: RMSD-Diagramm: Wildtyp(3LZM)-Mutante(3C80).....	58
Abbildung 49: RMSD-Diagramm: Wildtyp(3LZM)-Mutante(3C83).....	58
Abbildung 50: RMSD-Diagramm: Wildtyp(3LZM)-Mutante(3CDR).....	59
Abbildung 51: RMSD-Diagramm: Wildtyp(3LZM)-Mutante(3C8S).....	59
Abbildung 52: RMSD-Diagramm: Wildtyp(3LZM)-Mutante(3C7W).....	59
Abbildung 53: Energieprofilalignment: Wildtyp(2NTU)-Mutante(1P8I).....	60

Abbildung 54: Energieprofilalignment: Wildtyp(2NTU)-Mutante(1PXR).....	60
Abbildung 55: Energieprofilalignment: Wildtyp(2NTU)-Mutante(1PXS).....	60
Abbildung 56: Energieprofilalignment: Wildtyp(2NTU)-Mutante(1Q5I).....	61
Abbildung 57: Energieprofilalignment: Wildtyp(2NTU)-Mutante(1TN0).....	61
Abbildung 58: Energieprofilalignment: Wildtyp(2NTU)-Mutante(2I1X).....	61
Abbildung 59: Energieprofilalignment: Wildtyp(2NTU)-Mutante(2I21).....	62
Abbildung 60: Energieprofilalignment: Wildtyp(2NTU)-Mutante(2WJL).....	62
Abbildung 61: Energieprofilalignment: Wildtyp(2NTU)-Mutante(3HAQ).....	62
Abbildung 62: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3C7W).....	63
Abbildung 63: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3C7Y).....	63
Abbildung 64: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3C8R).....	63
Abbildung 65: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3C8S).....	64
Abbildung 66: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3C80).....	64
Abbildung 67: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3C83).....	64
Abbildung 68: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3CDQ).....	65
Abbildung 69: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3CDR).....	65
Abbildung 70: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3F8V).....	65
Abbildung 71: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(1P8I).....	66
Abbildung 72: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(1PXR).....	66
Abbildung 73: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(1PXS).....	66
Abbildung 74: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(1Q5I).....	67
Abbildung 75: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(1TN0).....	67
Abbildung 76: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(2I1X).....	67
Abbildung 77: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(2I21).....	68
Abbildung 78: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(2WJL).....	68
Abbildung 79: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(3HAQ).....	68
Abbildung 80: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3C7W).....	69
Abbildung 81: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3C7Y).....	69
Abbildung 82: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3C8R).....	69
Abbildung 83: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3C8S).....	70
Abbildung 84: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3C80).....	70
Abbildung 85: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3C83).....	70
Abbildung 86: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3CDQ).....	71
Abbildung 87: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3CDR).....	71

Abbildung 88: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3F8V).....	71
---	----

Tabellenverzeichnis

Tabelle 1: Die 20 kanonischen Aminosäuren.....	4
Tabelle 2: Überblick der Intermediate von Bacteriorhodopsin.....	10
Tabelle 3: Häufigkeit der einzelnen Aminosäuren nach ihrer Außen- bzw. Innenposition im Protein	21
Tabelle 4: Auswirkungen der Punktmutationen in Hinblick auf Stabilität, Struktur und Funktion (Bacteriorhodopsin).....	30
Tabelle 5: Auswirkungen der Punktmutationen in Hinblick auf Stabilität, Struktur und Funktion (Lysozym)	31
Tabelle 6: Q-Score-Tabelle von Bacteriorhodopsin berechnet mit dem Dali-Server.....	37
Tabelle 7: Q-Score-Tabelle von Lysozym berechnet mit dem Dali-Server.....	38
Tabelle 8: Anzahl der Basenänderungen in einem Triplet, um eine Mutation im Bacteriorhodopsin auszulösen.....	53
Tabelle 9: Anzahl der Basenänderungen in einem Triplet, um eine Mutation im T4 Lysozym auszulösen.....	54

Abkürzungsverzeichnis

BR	Bacteriorhodopsin
bzw.	beziehungsweise
DNS	Desoxyribonukleinsäure
eProS	Energy Profile Server
MEPAL	multiple energy profile alignment
PDB	Abkürzung für die vom RCSB betriebenen Protein Data Bank
PDBTM	Protein Data Bank of Transmembrane Proteins
RCSB	Research Collaboratory for Structural Bioinformatics
RMSD	root mean square deviation
TMDET	Detection of transmembrane regions
z.B.	zum Beispiel

Danksagung

Diese Bachelorarbeit wäre ohne die Unterstützung vieler Personen nicht möglich gewesen. Ich möchte mich bei meinem Hochschulbetreuer Prof. Dr. rer. nat. Dirk Labudde für die Unterstützung bei wissenschaftlichen Fragen und für die wertvollen Ratschläge bedanken. Ohne seine Anregungen wäre diese Arbeit nie auf Papier gelangt. Ebenfalls ein offenes Ohr hatte jederzeit Dipl.-Inf. (FH) Daniel Stockmann, dem ich an dieser Stelle für seine Hilfestellungen danken möchte. Weiterhin möchte ich meinen Dank Herrn Florian Heinke aussprechen.

Mein besonderer Dank gilt meiner Familie, die mich in meiner ganzen Studienzeit stark unterstützt und immer wieder aufgebaut hat. Ohne ihren Rückhalt in schwierigen Situationen wäre diese Abschlussarbeit nicht zustande gekommen. Auch bin ich ihnen sehr dankbar für die Suche nach Druckfehlern und Stilblüten. Ein großer Dank gilt auch meinem Freund, Steffen, ohne seine Unterstützung wäre die Zeit der Bachelorarbeit viel schwieriger gewesen.

1 Einleitung/ Motivation

Proteine besitzen auf der zellulären Ebene eine große Bedeutung. Sie bestehen aus Aminosäuren, deren Abfolge wesentlich zur Tertiär- und Quartärstruktur (stabile Struktur) des Proteins beitragen. Die Reihenfolge der Aminosäuren, auch Sequenz genannt, ist in der Desoxyribonukleinsäure (DNS) in jedem Protein gespeichert.

Somit sind Proteinmoleküle in der Lage sich selbst organisierend in 3D-Topologien zu falten, um so eine biologische Funktion erfüllen zu können. Wichtige Funktionen von Proteinen sind unter anderem die Aufrechterhaltung der Stabilität, Kommunikation innerhalb der Zelle, der aktive Stofftransport in der Zelle und das Binden oder katalytische Verändern von anderen Proteinen oder Molekülen. In dieser Arbeit werden globuläre Proteine und Membranproteine untersucht. Dabei sind die Membranproteine sehr wichtig. Sie dienen der Signalweiterleitung, den Pump- und Transportvorgängen und steuern wichtige zelluläre Funktionen. Ein bedeutender Repräsentant dieser Klasse ist das Bacteriorhodopsin, welches der Energiegewinnung im *Halobacterium salinarum* dient.¹ Wie genau sich aus einer Aminosäuresequenz eine Struktur bildet, ist noch schwer abzuschätzen. Der erste Ansatz für das Faltungsproblem stützt sich auf eine detaillierte strukturelle und energetische Beschreibung des Gleichgewichts zwischen dem nativen Zustand, den ungefalteten Proteinen und den zufällig ungeordneten Randzuständen. Schon durch ein paar Mutationen kann das Gleichgewicht zwischen gefalteten und ungefalteten Proteinen stark in beide Richtungen verschoben werden. Darum können schon wenige Mutationen neue tertiäre und/ oder sogar quartäre Strukturen hervorrufen. Eine geringe erbgutveränderte Störung kann eine völlig andere Struktur zur Folge haben und sich so auch auf die Leistungsfähigkeit des Proteins auswirken. Punktmutationen können dabei durch verschiedene Ursachen ausgelöst werden. Unterschiedliche Auslöser können sein: diverse Strahlen (Röntgenstrahlen, kosmische Strahlung, energiereiche Strahlen, radioaktive Strahlung), Fehler bei molekulargenetischen Prozessen (Replikation, DNA-Reparatur, Chromosomenver

¹ [Dressel, 2008]

teilung bei der Kernzellteilung), Viren und Schimmelpilze, sehr hohe bzw. niedrige Temperaturen, sowie chemische Substanzen (Senfgas, Formaldehyd). Es können auch Spontanmutationen auftreten, die ohne erkennbaren Grund in Erscheinung treten.² Es ist wichtig, die allgemeine Proteinfaltung zu kennen, um daraus neue Proteinstrukturen und -funktionen zu entwickeln. Die meisten natürlichen Proteine nehmen nur den nativen Zustand ein, mit einer freien Enthalpie von 5 kcal/mol. Es wird angenommen, dass manche Mutationen notwendig sind, um das Gleichgewicht von alternativen Zuständen auf eine freie Enthalpie von 5 kcal/mol zu verschieben. Diese Annahme stützt die meisten Methoden der Bioinformatik. Die überwiegende Anzahl der Mutationen steigern die Bereitschaft des Proteins zur Faltung in einen alternativen Zustand. So kann der Unterschied einer Sequenz von zwei Proteinen mit unterschiedlichen Strukturen sehr klein sein. Darum muss man die Proteinsequenzen methodisch untersuchen, um die Auswirkungen von Mutationen auf die Struktur und die Funktion derer ermitteln zu können.³

1.1 Ziel der Arbeit

Ziel dieser Arbeit ist es, die Auswirkungen von verschiedenen Punktmutationen auf zwei unterschiedliche Proteine zu analysieren. Besonderes Augenmerk liegt dabei auf der Stabilität, Funktion und Strukturänderung der Proteine. Die zentrale Frage ist, welchen Einfluss Punktmutationen auf die Stabilität und somit auch auf die Funktionalität der verschiedenen Proteine haben. Um dies herauszufinden, wurden vielfältige bioinformatische Tools verwendet. Insbesondere wurde mit dem Hochschulwebserver eProS⁴ gearbeitet.

2 [Pews, 2010]

3 [Alexander, 2009]

4 eProS ist ein Energy profile server der Hochschule Mittweida
(siehe <http://bioservices.hs-mittweida.de/Epros/>)

2 Biologische Grundlagen

2.1 Proteine

Proteine gehören zur Klasse der Makromoleküle einer Zelle und sind damit, neben den Polysacchariden, Lipiden und Nucleinsäuren, ein sehr wichtiger Bestandteil in einem Organismus. Sie besitzen eine strukturelle und enzymatische Bedeutung. So enthält eine Zelle tausende Proteine jeweils verschiedener Arten.

Ein Protein ist definiert als ein Polypeptid (oder eine Ansammlung von Polypeptiden), die ein Molekül mit spezifischer biologischer Funktion bilden. Diese Polypeptide bestehen aus Monomeren, auch Aminosäuren genannt. Es gibt 20 kanonische Aminosäuren (siehe Tabelle 1), die unterschiedliche Eigenschaften aufweisen. Sie besitzen dabei eine allgemeine Struktur (siehe Abbildung 1). Sie besteht aus einer Carboxylgruppe, einer Aminogruppe, einem α -Kohlenstoffatom und einem Rest, der je nach Aminosäure spezifisch ist. Dadurch können sich die Aminosäuren zu einer Peptidbindung verknüpfen.⁵ So entsteht mit Hilfe der Ribosomen bei der Translation eine Aminosäurekette, die wiederum als Protein bezeichnet wird. Diese Proteine können strukturell in zwei Klassen eingeteilt werden, den globulären Proteinen und den Membranproteinen.

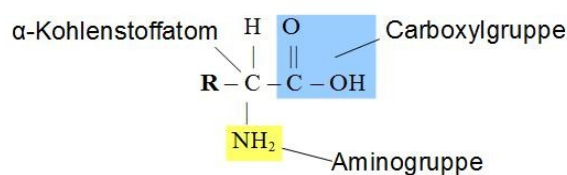


Abbildung 1: Allgemeine Struktur einer Aminosäure

5 [Madigan, 2006]

Tabelle 1: Die 20 kanonischen Aminosäuren

Aminosäure	Dreibuchstaben-code	Einbuchstaben-code	Durchschnittlicher Vorkommen in Proteine	Polarität von R
Glycin	Gly	G	7,2 %	unpolar
Alanin	Ala	A	7,8 %	unpolar
Valin	Val	V	6,6 %	unpolar
Leucin	Leu	L	9,1 %	unpolar
Isoleucin	Ile	I	5,3 %	unpolar
Methionin	Met	M	2,3 %	unpolar
Phenylalanin	Phe	F	3,9 %	unpolar
Tryptophan	Trp	W	1,4 %	unpolar
Prolin	Pro	P	5,2 %	unpolar
Serin	Ser	S	6,8 %	polar
Threonine	Thr	T	5,9 %	polar
Asparagin	Asn	N	4,3 %	polar
Glutamin	Gln	Q	4,2 %	polar
Cystein	Cys	C	1,9 %	polar
Tyrosin	Tyr	Y	3,2 %	polar
Aspartat	Asp	D	5,3 %	polar
Glutamat	Glu	E	6,3 %	polar
Lysin	Lys	K	5,9 %	polar
Arginin	Arg	R	5,1 %	polar
Histidin	His	H	2,3 %	polar

2.1.1 Membranproteine

Membranproteine sind in eine dünne Membran eingelagert, die das Cytoplasma einer Zelle umgibt. Sie gewährleistet die Konstanz der Zellinhaltsstoffe gegenüber dem umgebenden Milieu, ihre Permeabilität ermöglicht die gezielte Aufnahme und Abgabe von Verbindungen. Sie dient zur Speicherung von Zellenergie. Die Membranproteine sind spezielle Proteine, die eine Kommunikation und den Stoffaustausch mit der Umgebung ermöglichen. Weiterhin nützen sie dem Aufbau von Zellenergie. Anders als globuläre Proteine exponieren Membranproteine die Seitenketten hydrophober Aminosäuren auf ihre Oberfläche. Dabei unterscheidet man zwischen integrale und periphere Membranproteine.⁶

Auf der einen Seite gibt es die integralen Membranproteine (siehe Abbildung 2), die bestimmte Moleküle oder Ionen (z.B. Natrium, Kalium, Kalzium oder Chlorid) als einfache Porenproteine oder Ionenkanäle durch die Membran lassen. So können mit Hilfe von Carrierproteinen bestimmte Stoffe sehr spezifisch von innen nach außen oder umgekehrt transportiert werden. Dies ist auch gegen den Konzentrationsgradienten möglich, wobei Energie verbraucht wird. Unter die integralen Proteine fallen auch die Rezeptorproteine. Sie können mit Hormonen und Neurotransmittern wechselwirken und somit einen Informationsaustausch zwischen dem Zelläußeren und -inneren bewirken. Außerdem besitzen sie mechanische Funktionen. So verknüpfen sie die Zelle mit einer Nachbarzelle oder mit bestimmten Proteinen.⁷

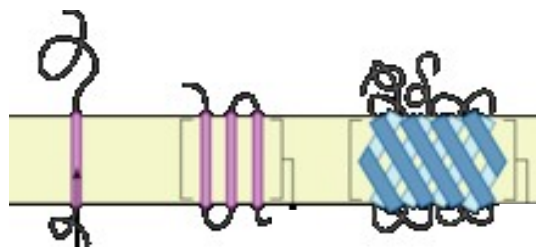


Abbildung 2: Integrale Membranproteine⁸

6 [Helmich, 2011]

7 [Deplanque, 2011]

8 [Foobar, 2006]

Auf der anderen Seite gibt es periphere Membranproteine, die in Lipid-Einzel-schichten eingelagert sind oder sich liegend auf der Membran befinden.⁶ Sie können sich jedoch auch an integrale Proteine binden (siehe Abbildung 3). Durch hohe Salzkonzentrationen oder durch Veränderung des pH-Wertes kann man die peripheren Membranproteine leicht von der Membran lösen.⁷

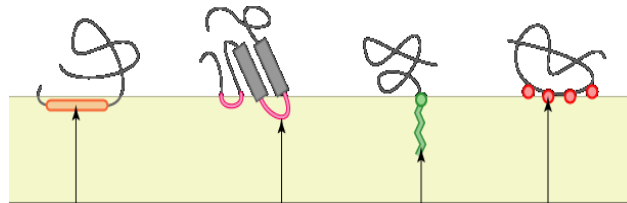


Abbildung 3: Periphere Membranproteine⁸

2.1.1.2 Bacteriorhodopsin

Die Photosynthese als Energiegewinnung ist direkt oder indirekt für alle Lebewesen sehr wichtig. Die Entdeckung des Proteins Bacteriorhodopsin in den 70er Jahren⁹ verwarf das Bild, dass die Umwandlung von Lichtenergie in chemische Energie nur über das Chromophor Chlorophyll abläuft. Bacteriorhodopsin stellt einen anderen Photosynthesemechanismus dar.

In dieser Arbeit wurde das integrale Membranprotein Bacteriorhodopsin (siehe Abbildung 5) gewählt. Es ist ein stäbchenförmiges, vier Flagellen tragendes prokaryotisches Archaeobakterium, dass im *Halobacterium salinarium* vorkommt.¹⁰ Da dieser Organismus in einem extremen Lebensraum mit hohen Salzkonzentrationen und Temperaturen beheimatet ist, musste er sich mit einer

6 [Helmich, 2011]

7 [Deplanque, 2011]

8 [Foobar, 2006]

9 [Oesterhelt, 1971]

10 [Kandt, 2003]

alternativen Energiegewinnung anpassen. Da der Sauerstoffgehalt mit höher werdender Temperatur und Salzgehalt abnimmt, benutzt Bacteriorhodopsin keine sauerstoffabhängigen Prozess für das Erzeugen eines Protonengradienten. Der Protonenkonzentrationsgradient entsteht mit Hilfe von Lichtenergie (lichtgetriebene Protonenpumpe), dabei fungiert Bacteriorhodopsin in der Zellmembran als einfaches Photosynthesesystem.¹¹ Die Produktion von ATP wird sichergestellt, indem Membransegmente (Purpurmembranen) entstehen und somit eine Erhöhung der Synthese stattfindet. Das Protein bildet in der Purpurmembran ein zweidimensionales Gitter aus (siehe Abbildung 4).

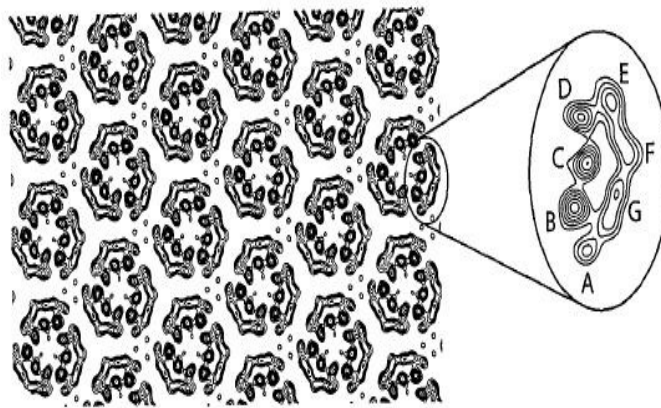


Abbildung 4: Karte der Elektronendichte eines zweidimensionalen Bacteriorhodopsin-Gitters.¹¹

Man erkennt die hexagonale Anordnung der Proteintrimere in der Membran und den innermolekularen Aufbau mit den sieben senkrecht zur Membran stehenden Helices A bis G.

Bacteriorhodopsin hat ein Molekulargewicht von 26,8 kDa und besteht aus 249 Aminosäuren und einem Retinalmolekül, welches an Lysin 216 mittels einer Schiff'schen Base gebunden ist. Die Aminosäuren falten sich in sieben Helices, die das im Innern liegende Retinal einschließen (siehe Abbildung 5). Im Grundzustand liegt das auch in Sehzellen vorkommende Molekül in all-trans-Konfor

¹¹ [Oesterhelt I, 1971]

mation vor und bei Lichtabsorption findet eine all-trans/13-cis-Isomerisierung statt. Durch das Abwechseln dieser zwei Zustände, kommt es zum Funktionszyklus, während dessen ein Proton durch das Molekül hindurch auf die andere Seite der Membran transportiert wird. Das Protein besitzt eine cytoplasmatische und eine extrazelluläre Seite, die jeweils für die Protonenaufnahme bzw. -abgabe zuständig sind. Diese werden durch das Retinal voneinander getrennt. Wichtige Elemente für das Durchschleusen der Protonen im Molekül sind die Aspartate und Glutamate, die eine Bindung des Protons mit Bindungsenergien im Bereich der Bildung von Hydroniumionen ermöglichen. Diese stehen im Wettbewerb mit den Wassermolekülen, um die Anlagerung von Protonen und können durch ein Anheben oder Absenken des pK-Wertes¹² der Gruppe geschaltet werden. Da jedes aufgenommene Proton auch wieder abgegeben werden muss, dürfen die Unterschiede der pK-Werte nicht zu groß sein.¹¹



Abbildung 5: Struktur von Bacteriorhodopsin¹³

12 Der pK-Wert ist der negative dekadische Logarithmus der Gleichgewichtskonstanten K einer chemischen Reaktion.

11 [Bolwien, 2002]

13 [Cronk, 2011]

Der Photozyklus von Bacteriorhodopsin ist in Intermediate klassifiziert (siehe Abbildung 7), da das Retinal abhängig von der Zeit seine Farbe ändert. Einen Überblick dieser Intermediate mit deren Veränderungen ist in Tabelle 2 zusammengefasst. In Abbildung 6 ist der Protonenweg innerhalb des Proteins zu erkennen. Der untere Bereich (zuständig für die Protonenabgabe) bildet einen Protonenleitweg. Angefangen bei der Schiff'schen Base, über Asp 85, ein Wassermolekül zwischen Asp 85 und Asp 212 und endet an der Dyade Glu 194/Glu 204, die in unmittelbarer Nähe der Oberfläche des Proteins liegen. Zwischen Asp 85 und der Dyade liegt Arg 82, dem eine schaltende Funktion angedacht wird.¹¹

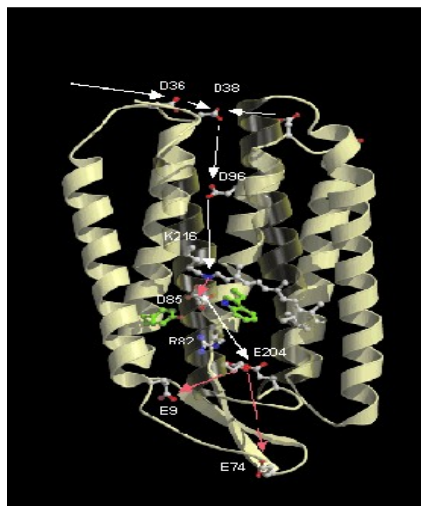


Abbildung 6: Protonenweg innerhalb von Bacteriorhodopsin¹⁴

Der dreidimensionale Aufbau von BR im Grundzustand in einer vereinfachten Darstellung, bei der nur das Proteinrückgrat, das Retinal und wichtige Aminosäurenreste dargestellt sind. Die Glutamate und Aspartate im unteren, extrazellulären Bereich des Proteins bilden bereits im Grundzustand eine Kette von Wasserstoffbrückenbindungen, die den Weg des Protons vorzeichnen. Im cytoplasmatischen Teil des Proteins sind zur Reprotonierung der Schiff'schen Base Strukturänderungen nötig. Die entstehenden Hohlräume können Wassermoleküle aufnehmen, die dann als Wasserstoffbrückennetzwerk das Verschieben eines Protons erlauben.

¹¹ [Bolwien, 2002]

¹⁴ [Essen, 1998]

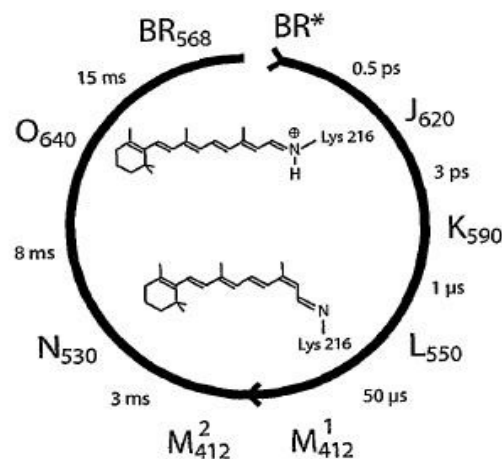


Abbildung 7: Photozyklus von Bacteriorhodopsin¹¹

Photozyklus, wie er anhand der Farbänderungen des Retinals festgestellt werden kann. Lediglich die Einteilung in die beiden M-Zustände erfolgt aufgrund mechanistischer Überlegungen und beschreibt die Änderung der Zugänglichkeit der Schiff'schen Base.

Tabelle 2: Überblick der Intermediate von Bacteriorhodopsin¹¹

Intermediate	Veränderung
$BR_{568} \rightarrow J_{620}$	Das Retinal befindet sich im lichtadaptierten Grundzustand, der Absorbierung eines Photons folgt die Isomerisierung um die C13 = C14 Doppelbindung.
$J_{620} \rightarrow K_{590} \rightarrow L_{550}$	Die im Retinal gespeicherte Energie dissipiert ¹⁵ und sorgt für erste, relativ geringe Änderungen im sichtbaren Spektrum.
$L_{550} \rightarrow M_{412}$	Das Proton der Schiff'schen Base verschiebt sich zu Asp 85. Gleichzeitig wird ein Proton in das extrazelluläre Medium abgegeben. Die aus der Abgabe des Protons stark veränderte Elektronenstruktur des Retinals ist an der starken Blauverschiebung des Absorptionsmaxi-

¹⁵ Bei allen dissipativen Prozessen wird Entropie produziert, wobei die dissipierte Energie auf die produzierte Entropie umgeladen wird.

¹¹ [Bolwien, 2002]

	mums erkennbar.
$M_{412} \rightarrow M'_{412}$	Die Zugänglichkeit der Schiff'schen Base ändert sich zugunsten der cytoplasmatischen Seite, um eine Reprotonierung zu ermöglichen.
$M'_{412} \rightarrow N_{530}$	Die Schiff'sche Base wird durch Asp 96 reprotoniert.
$N_{530} \rightarrow O_{640}$	Asp 96 wird aus Richtung des Cytoplasmas reprotoniert und ein Proton aus dem Inneren der Zelle wird aufgenommen. Gleichzeitig findet eine Rückkehr des Retinals in den all-trans-Grundzustand statt.
$O_{640} \rightarrow BR_{568}$	Asp 85 gibt sein Proton ab und das Protein kehrt damit in den Grundzustand zurück.

2.1.2 Globuläre Proteine

Globuläre Proteine sind sphäriode Moleküle, wie zum Beispiel Enzyme, Transportproteine und Rezeptoren. Sie können sowohl alpha-Helices als auch beta-Faltblattstrukturen enthalten (durchschnittlich ca. 27% alpha-Helix und 23% beta-Faltblatt). Die Seitenketten sind je nach Polarität anders angeordnet. Unpolare Reste, wie Val, Leu, Ile, Met, Phe, befinden sich vorwiegend im Inneren des Proteins. Anders verhält es sich mit geladenen (Arg, His, Lys, Asp, Glu) und ungeladenen Resten (Ser, Thr, Asn, Gln, Tyr, Trp), die zumeist an der Oberfläche sind.¹⁶ Ihre Polypeptidketten sind nahezu kugelförmig zusammen gelagert. Globuläre Proteine sind wichtige Bestandteile des Protoplasmas der Zellen. Daher haben sie nicht nur als Nährstoffe, sondern auch als Reservestoffe für den Baustoffwechsel Bedeutung. Sie kommen im Blut und in den Muskeln, sowie als Bestandteile von Enzymen und Hormonen vor.¹⁷

¹⁶ [Albrecht, 2011]

¹⁷ [Liebbert, 1986]

2.1.2.1 T4 Lysozym

Das T4 Lysozym ist ein kleines globuläres Protein mit 164 Aminosäuren und 1429 Atome. Das Molekulargewicht beträgt 19 kDa. Strukturell aufgeklärt wurde es mittels Kristallstrukturanalyse, NMR-Spektroskopie und ESR-Spektroskopie. Man kann zwei Domänen unterscheiden. Die N-terminale Domäne enthält das einzige β -Faltblatt und den größten Teil der flexiblen Elemente. Sie ist durch eine lange α -Helix mit der C-terminalen Domäne verbunden, die aus einem Barrel von sieben α -Helices besteht (siehe Abbildung 8) und in der sich die meisten der positiven Überschussladungen befinden. Das T4 Lysozym ist ein basisches Protein, welches einen Isoelektrischen Punkt von über 9 besitzt und neunfach positiv geladen ist, wenn es sich in einer neutralen pH-Wert-Umgebung befindet.

Lysozyme aus dem Bakteriophagen T4 gehören zur der Klasse der Enzyme. Die biologische Funktion der T4-Lysozyme liegt in der Spaltung von Peptidoglykane der Zellwände des Wirtsbakteriums.¹⁸ Somit lysiert die Zelle und die Phagen werden freigesetzt.

Das aktive Zentrum befindet sich zwischen den beiden Domänen und bindet ein Substrat, in dem es eine Art Scharnierbewegung vollzieht. T4-Lysozym Mutanten zeigen häufig gleiche Kristallbedingungen, können jedoch die Funktion und Stabilität des Proteins beeinflussen.¹⁹

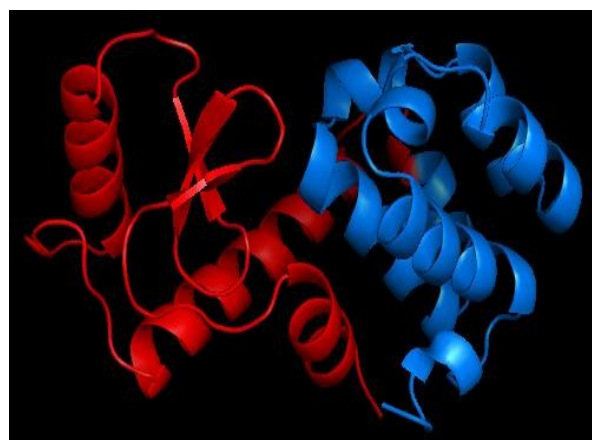


Abbildung 8: Tertiärstruktur des T4 Lysozym Proteins

18 [Jacobsen, 2005]

19 [Böhm, 2005]

2.2 Wechselwirkungskräfte in Proteinen

Damit sich ein Protein falten kann, benötigt es verschiedenartige Wechselwirkungskräfte. Um die korrekte Faltung anzunehmen (und nicht eine der „unendlich“ vielen denaturierten Konformationen), müssen die Wechselwirkungskräfte den entropisch sehr ungünstigen Zustand kompensieren.

2.2.1 Kovalente Bindungen

Kovalente Wechselwirkungen sind sehr starke Kräfte, bei denen sich die Atome zu mehr oder weniger gleichen Teilen die Elektronen teilen. Sie weisen eine Reichweite von 0,1 nm bis 0,15 nm auf. Die Veränderung der chemischen Natur ist jedoch immer noch möglich. Durch diese Art von Bindung bilden sich Molekülgefüge, die keine selbstständigen Einheiten mehr darstellen.²⁰ In Abbildung 9 ist am Beispiel von zwei Chloratomen die Bildung einer kovalenten Bindung veranschaulicht. Diese Bindung wird auch Molekülbindung genannt.²¹

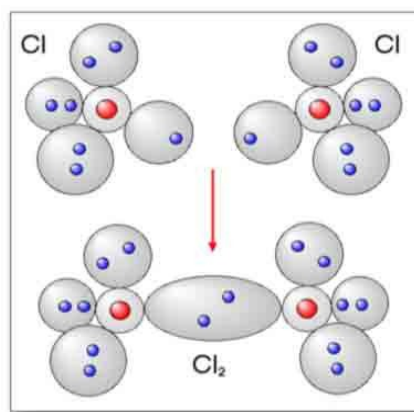


Abbildung 9: Kovalente Bindung zweier Chloratome²¹

²⁰ [Brumm, 2010]

²¹ [Domain, 2011]

In einem Protein treten derartige Bindungen als Peptidbindungen auf, die die einzelnen Aminosäuren miteinander verknüpfen. Peptidbindungen sind durch ihre Mesomerie stabilisiert. Sie bedingt den partiellen Doppelbindungscharakter der Amidgruppe, was dazu führt, dass die Peptidbindung relativ starr und nicht drehbar ist.²² Die amidische Bindung im Molekül ist eine planare Bindung, die eine ebene Fläche bildet. Der partielle Doppelbindungscharakter äußert sich auch im Bindungsabstand zwischen dem C- und N-Atom, der in der Nähe des Bindungsabstands einer C=N-Doppelbindung, ca. bei 0,124 nm (C-N-Einfachbindung: 0,147 nm), liegt. Die Bindung zwischen dem α -Kohlenstoffatom und dem Carbonylkohlenstoffatom ist eine reine Einfachbindung, genau wie jene zwischen dem α -Kohlenstoffatom und dem Stickstoffatom. Folglich weisen diese Bindungen eine große Rotationsfreiheit auf.²³ Der Rotationswinkel wird vor allem von sterischen Einschränkungen beeinflusst.

2.2.2 Dipol-Dipol-Wechselwirkungskräfte

Voraussetzung für eine Dipol-Dipol-Bindung ist eine starke Elektronegativitätsdifferenz in einem Molekül. Durch eine permanente Ladungsverschiebung entstehen Dipole, die sich gegenseitig elektrostatisch anziehen, wobei jedoch kein Ladungsaustausch stattfindet. Die Dipol-Dipol-Kräfte sind um so größer, je polarer die Moleküle gebaut sind. Außerdem wirken sie bei polaren Molekülen zusätzlich zu den immer vorhandenen Van-der-Waals-Kräften.

2.2.3 Wasserstoffbrückenbindungen

Eine besondere Art der Dipol-Dipol-Bindung ist die Wasserstoffbrückenbindung. Eine Wasserstoffbrückenbindung liegt dann vor, wenn das Wasserstoffatom mit einem oder mehreren Atomen verbunden ist (siehe Abbildung 10). Dabei dient ein H-Atom als Brücke zwischen zwei stark elektronegativen Atomen, wobei es mit

22 [Koolman, 1998]

23 [Voet, 2002]

einem Atom durch eine polare kovalente Bindung und mit dem anderen (z.B. Stickstoff-, Sauerstoff- oder Fluoratom) durch elektrostatische Kräfte verbunden ist. Wasserstoffbrückenbindungen sind äußerst wichtig für die Strukturbildung und der Stabilität von Proteinen. Die Stärke der Bindung beläuft sich auf ca. 20 kJ/mol. Die Anziehung ist somit viel geringer als die kovalente Bindung, die eine Bindungsenergie von ca. 200 bis 400 kJ/mol aufweist.²⁴

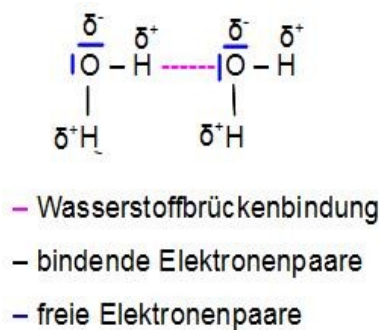


Abbildung 10: Wasserstoffbrückenbindung

2.2.4 Van-der-Waals-Kräfte

Auch zwischen völlig unpolaren Molekülen und Atomen gibt es Wechselwirkungen. Van-der-Waals-Kräfte bezeichnen relativ schwache nicht-kovalente Bindungen zwischen Atomen bzw. Molekülen. Diese beruhen auf schnellen und kurzen Schwankungen der Elektronenverteilung in Atomen und Molekülen. In dieser kurzen Zeit wirken die Teilchen ähnlich wie Dipole. Die Kräfte sind sehr bedeutend für die Proteinfaltung und bei der Protein-Protein-Interaktion.²⁰

²⁴ [Fischer, 2010]

²⁰ [Brumm, 2010]

2.2.5 Abstoßungskräfte

Abstoßungskräfte sind sterische Wechselwirkungen, die nur eine geringe Reichweite besitzen. Kommen sich jedoch zwei Atome näher, vergrößert sich die Stärke der Kraft erheblich. Atomverbände reagieren deutlich mit Abstoßungsreaktionen, wenn man sie verdichten will. Ähnlich verhalten sich Elektronenwolken, die sich gegen eine zu große Annäherung sperren. Werden Atome starkem Druck ausgesetzt, ziehen sich die Elektronenwolken enger um einen Kernbereich zusammen.²⁵ Der Zustand ist in Hinblick auf die Energie sehr ungünstig, weswegen dieser in der Realität nicht vorkommt.

2.2.6 Hydrophober Effekt

Bei der Proteinfaltung ist der hydrophobe Effekt sehr wichtig. Wenn in der Umgebung von vorhandenen Wassermolekülen ein apolares Molekül oder eine apolare Region eines Moleküls vorliegt, unterliegen die Wassermoleküle einer Art Selbstorganisation. So befinden sich die Wassermoleküle in einem höheren Energiezustand, als wenn sie nur von Wassermolekülen umgeben wären. Wenn sich nun zwei apolare Gruppen annähern, die von geordneten Wassermolekülen umgeben sind, so gehen diese Wassermoleküle in einen ungeordneten Zustand über. Es resultierte ein Entropieverlust und somit auch die Abnahme der Freien Energie, was zur Folge hat, dass das Protein stabilisiert wird. Dabei ist die Hydrophobizität von Aminosäuren direkt von ihrer Oberfläche abhängig (siehe Abbildung 11).²⁶

25 [Maurer, 2006]

26 [Biochem, 2011]

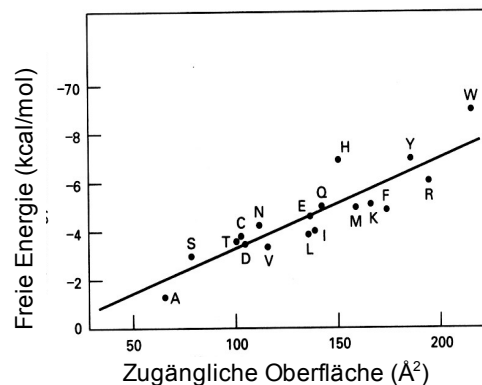


Abbildung 11: Abhängigkeit der Hydrophobizität der Aminosäuren und ihrer Oberfläche²⁶

2.3 Proteinfaltung

Über die Proteinfaltung ist bisher noch wenig bekannt. Doch besitzt sie eine große Bedeutung für die Biologie, da durch fehlerhaft gefaltete Proteine Krankheiten ausgelöst werden können. Der korrekte Faltungsablauf geschieht in sehr kurzer Zeit (einer Milliardstelsekunde (10^{-9}) bis zu Sekunden). Dabei wechselt das Protein vom denaturierten (unstrukturierten oder biologisch nicht funktionsfähigen) Zustand zum nativen (strukturierten oder biologisch funktionsfähigen) Zustand. Darum muss die sogenannte Energielandschaft die Faltungswege und die Geschwindigkeit bestimmen. Man geht davon aus, dass die Energielandschaft relativ übersichtlich ist und an einer definierten Stelle ein Optimum besitzt. Das Protein findet dieses Optimum instinktiv. Damit ein Protein stabil gegenüber thermodynamischen Fluktuationen ist, sollte der native Zustand eine möglichst tiefe freie Energie besitzen oder zumindest in einem tiefen Tal der Energielandschaft liegen. Dies ist Voraussetzung dafür, dass das Protein seine Funktion fehlerfrei ausüben kann. Dabei wird die räumliche Anordnung des Proteins von der energetischen Situation bestimmt.²⁷ Da ein Protein eine große Anzahl von möglichen Konformationen besitzt, müssen diese zu Beginn der Faltung begrenzt werden. Dies erfolgt durch die Van-der-Waals-Kräfte, die ein

²⁵ [Maurer, 2006]

²⁷ [Wachtveitl, 2011]

schnelles Zusammenlagern der hydrophoben Aminosäuren im polaren Cytosol bewirken. So wird aus der random-coil-Struktur ein kompaktes ungeordnetes Knäuel, welches auch random-globule-Zustand genannt wird (siehe Abbildung 12). Mit diesem Knäuel beginnt die eigentliche Faltung. Diese beiden Zustände sind durch eine definierte Barriere der Freien Energie voneinander getrennt.²⁸

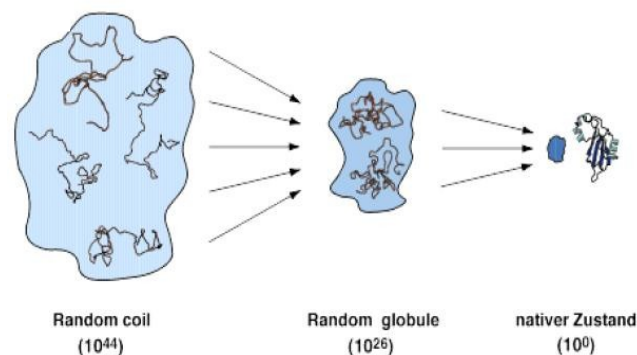


Abbildung 12: Schematische Darstellung der Proteinfaltung. Die Zahlen in den Klammern geben die Anzahl der Konformationen an.²⁷

Das ungefaltete Protein kann sehr leicht mit anderen Ketten Aggregate bilden, da es sich in einer hohen Konzentration im Cytosol befindet. Diese Gebilde können sich nun nicht mehr richtig falten. Eine weitere Gefahr geht von den einzelnen Aminosäureketten aus, die eine Missfaltung durch ihre zugehörigen lokalen Energieminima nicht mehr beheben können. Das Protein im ungefalteten Zustand kann jedoch auch sehr leicht von Proteasen angegriffen werden. Damit diese Probleme nicht auftreten, unterstützen Chaperone die Proteine bei der Faltung. Diese schirmen hydrophobe Regionen ab und verhindern somit eventuelle Aggregationen.²⁹

²⁷ [Wachtveitl, 2011]

²⁸ [Dobson, 1999]

²⁹ [Behre, 2002]

3 Bioinformatische Grundlagen

3.1 Biodatenbanken

Datenbanken besitzen in der Bioinformatik eine große Bedeutung. Da sich die Bioinformatik mit der Untersuchung der Molekularbiologie beschäftigt und dabei sehr große Datenmengen anfallen (und diese eine umfangreiche Datenanalyse benötigen), sind Datenbanken eine effektive Methode, diese zu hinterlegen und wieder auf sie zugreifen zu können.

Bei der vorliegenden Arbeit werden nur Proteinsequenz- und Proteinstruktur-Datenbanken verwendet. Es gibt jedoch auch andere Biodatenbanken. Unter anderen Sequenz-Datenbanken zur DNA-Analyse und Sequenzierung (zum Beispiel: GenBank) und Sequenz-Datenbanken zur Ermittlung phylogenetischer Bäume.³⁰

Die benötigten Informationen über das Bacteriorhodopsin Protein und das T4 Lysozym Protein wurden der Protein Data Bank (kurz PDB) entnommen. In dieser Datenbank sind die für diese Arbeit relevanten, dreidimensionalen Daten gespeichert. Die PDB ist weltweit ein einzigartiges Archiv für Strukturdaten von biologischen Makromolekülen. Sie sind für jedermann frei zugänglich. Dabei haben Wissenschaftler aus der ganzen Welt die hoch qualitativen Daten durch Experimente ermittelt und für die Datenbank bereitgestellt.³¹

Somit ist die Basis der Proteinanalyse geschaffen, auf der strukturellen Ebene sowie auch für die energiebasierende Analyse. Die Strukturdaten wurden üblicherweise durch Kristallstrukturanalyse oder NMR-Spektroskopie ermittelt. Die Mutanten der Proteine wurden ebenfalls mit Hilfe der PDB gefunden, wohingegen die Angaben über die Auswirkungen der Mutationen aus der Literaturdatenbank PubMed, die biomedizinische Literatur aus MEDLINE, Life-Science-Zeitschriften und Online-Bücher beinhaltet, herausgelesen wurden.

30 [François, 2002]

31 [Schmieder, 2004]

3.2 Energieprofile

Die Analyse von Proteinen anhand von Energieprofilen ist noch weitestgehend unerforscht, da bisher noch keine derartigen Einträge in Proteindatenbanken vorliegen. Deshalb müssen die Energiewerte für die einzelnen Aminosäuren berechnet werden.

Ein Energieprofil stellt dabei eine Umwandlung einer dreidimensionalen Struktur in einen zweidimensionalen Vektor dar. Wenn man die Energien für jede einzelne Aminosäure über eine Sequenz aufträgt, erhält man ein Energieprofil. Dafür wird ein Algorithmus von Frank Dressel benutzt, den er in seiner Dissertation „Sequenz, Energie, Struktur - Untersuchungen zur Beziehung zwischen Primär- und Tertiärstruktur in globulären und Membran-Proteinen“ vorstellt. Das Folgende wird seiner Arbeit entnommen.¹

Um ein Energieprofil zu berechnen, benötigt man die Gesamtenergie eines Proteins. Diese ergibt sich aus den Wechselwirkungsenergien der einzelnen Aminosäuren.

$$E_{Ges} = \sum_{ij} e_{ij}^{\circ} f(r_{ij}) + \sum_i e'_{i0} g(i) \quad (1)$$

e'_{i0} beschreibt dabei die Wechselwirkung der Aminosäuren mit dem umgebenden Lösungsmittel und e_{ij}° gibt die Wechselwirkungen von zwei Aminosäuren untereinander wider. Weiterhin beschreibt $f(r_{ij})$ eine Abstandsfunktion zwischen der Aminosäure i und der Aminosäure j und $g(i)$ den Zustand der Aminosäure i . Mit Hilfe einer Wahrscheinlichkeitsverteilung wird jeder Aminosäure eine Löslichkeitsenergie e_i zugewiesen. Dabei wird unterschieden, ob die Aminosäure hydrophil ist, das heißt, dass sie dem Lösungsmittel zugewandt ist und mit diesem Wasserstoffbrückenbindungen ausbildet, oder im Protein innen angelagert ist (hydrophob) und somit mit anderen Aminosäuren interagiert. $g(i)$ beschreibt also ein Kriterium der Aminosäure (innen/außen). Um herauszufinden, ob eine Aminosäure im Inneren eines Proteins liegt oder außen, wurden

¹ [Dressel, 2008]

zwei Kriterien formelmäßig erfasst. Man geht davon aus, dass ein Zentrum (der Schwerpunkt) c der C_α -Atome definiert ist und um diese eine Kugel gespannt wird, mit dem Radius von 10 Å, um den Innen-/Außen-Zustandsraum zu definieren. Die Innen-/Außenverteilung der einzelnen Aminosäuren ist in Tabelle 3 aufgeführt. Wenn nun $g(i)=1$ und die Bedingungen der Gleichung (2) erfüllt sind, gilt die Aminosäure als innen liegend. Ist die Bedingung nicht erfüllt und $g(i)=0$, so liegt die Aminosäure außen.

$$|C_\alpha - c| < 5 \vee (C_\alpha - C_\beta)(C_\alpha - c) < 0 \quad (2)$$

Tabelle 3: Häufigkeit der einzelnen Aminosäuren nach ihrer Außen- bzw. Innenposition im Protein¹

Aminosäure	innen	außen	Aminosäure	innen	außen
Cys	4582	1016	His	6419	3366
Ile	20370	4141	Gly	16698	14326
Ser	12576	10411	Asp	10001	14327
Gln	7373	7752	Leu	30615	7107
Lys	9285	15193	Arg	11327	10441
Asn	8225	8928	Trp	4001	1193
Pro	9135	9423	Val	23562	6551
Thr	12537	9622	Glu	11165	18091
Phe	13353	2813	Tyr	11228	3529
Ala	22725	11052	Met	7003	1723

Die Energiedifferenz zwischen dem hydrophilen und hydrophoben Zustand kann man nun mit der Boltzmannverteilung berechnen:³²

¹ [Dressel, 2008]

³² [Schneider, 2009]

$$e'_{i0} = -k_B T \ln \left(\frac{n_{in}}{n_{out}} \right) \quad (3)$$

n_{in}/n_{out} bezeichnet hierbei die Anzahl der Fälle, in dem die betrachtete Aminosäure im Protein oder auf der Oberfläche vom Protein liegt.¹ Dabei wird eine Statistik aus verwendeten Strukturdaten benötigt.

Die Boltzmannsche Konstante und die Temperatur T werden dabei als konstant angesehen. Um einen Abstand zwischen zwei Aminosäuren berechnen zu können, muss man einen räumlichen Anhaltspunkt schaffen (dies wird ebenfalls zur Berechnung von r_{ij} benötigt). Da die Seitenketten der Aminosäuren durch ihre Eigenschaften in ihrer Dynamik sehr unterschiedlich sind, werden die C_β -Atome zur Orientierung gewählt, da diese relativ stark an der Backbone gebunden sind. Die aus der Formel (3) errechneten Energiewerte geben die Energie wieder, die aufgewendet werden muss, um eine Aminosäure vom hydrophilen Zustand in den hydrophoben Zustand zu überführen und andersrum. Wenn man nun diese Werte über eine Sequenz aufträgt, erhält man ein Löslichkeitsenergieprofil.

Ein anderer Ansatz ist der, dass sich die Wechselwirkungsenergien zwischen zwei Aminosäuren von den beobachteten und statistischen Kontakten ableiten lassen. Alle Kontakte von Aminosäuren i und j wurden ausgezählt und als n_{ij} zusammengefasst. $N_{contact}$ gibt die Anzahl der Kontakte an. Zusätzlich mit den relativen Häufigkeiten p_i und p_j lässt sich folgende Gleichung aufstellen:

$$e^\circ_{ij} = -k_B T \ln \left(\frac{n_{ij}}{N_{contact} p_i p_j} \right) \quad (4)$$

¹ [Dressel, 2008]

Da man aber das reine Paarpotential möchte, muss man an der Gleichung (3) noch einige Änderungen vornehmen:

$$\mathbf{e}_{io} = \left(\frac{1}{i} \alpha_i \right) \mathbf{e}'_{io} \quad (5)$$

α_i gibt die Anzahl der Kontakte an, die die Aminosäure i mit den Aminosäuren j hat. Ein Kontakt ist definiert durch $r_{ij} < 8 \text{ \AA}$. Das einfache Paarpotential lässt sich wie folgt formulieren:

$$\mathbf{e}_{ij} = \mathbf{e}_{io} + \mathbf{e}_{jo} + \mathbf{e}_{ij}^{\circ} \quad (6)$$

Um ein Energieprofil erstellen zu können, wird nun für jede Aminosäure im Protein die Energie berechnet und über ihre Primärstruktur aufgetragen. Ein Energieprofil ist für jedes Protein einzigartig und kann somit zur Untersuchung von Mutationen gut eingesetzt werden. Man benötigt jedoch eine dreidimensionale Struktur des Proteins, da man die Koordinaten zur Berechnung und Darstellung braucht. In vorhergehenden Arbeiten wurde dieser Algorithmus implementiert und auf dem Server eProS (siehe <http://bioservices.hs-mittweida.de/Epros/>) der Hochschule Mittweida der Öffentlichkeit zur Verfügung gestellt. Jetzt bedarf es nur noch einer Datei in PDB-Format oder der PDB-ID, die als Eingabe dient. Denn alle relevanten Informationen (Sequenzinformationen, Koordinaten), die für die Berechnung eines Energieprofils verwendet werden, beinhalten diese Dateien.

3.3 Tools für die Mutationsanalyse

3.3.1 DaliLite

DaliLite ist ein Programm für den paarweisen Vergleich zweier Strukturen, aber dient auch der Strukturdatenbanksuche. Es ist eine eigenständige Version der beliebten Suchmaschine des Dali-Servers. Dieses Programm trägt dazu bei, dass Forscher eine große Anzahl von Strukturen effizient und lokal vergleichen können. Datenbasis sind die räumlichen Abstände der C α -Atome, die für jede Kombination von Residuen einer Struktur in eine Distanzmatrix eingetragen werden. Dabei ergibt sich stets der selbe Matrizeninhalt für eine Struktur, wobei die Rotations- und Translationsoperationen keinen Einfluss haben. Schließlich wird die gewählte Kombination als strukturelles Alignment ausgegeben.³³

Ausführlicher ist der Dali-Algorithmus ein Superpositionier-Verfahren, welches von C. Sander und L. Holm entwickelt wurde. Er vergleicht Proteinstrukturen mit zweidimensionalen Matrizen, wobei jedes Element in der Matrix die euklidische Distanz zwischen dem i-ten und j-ten Rest des Proteins darstellt. Die Distanzmatrizen werden paarweise align und die n homologen Reste identifiziert. Die strukturelle Ähnlichkeit der beiden Proteine (A und B) wird dann wie folgt berechnet:

$$S = \sum_i \sum_j \left(0,2 - \frac{|d_{ij}^A - d_{ij}^B|}{d_{ij}^o} \right) e^{-\left(\frac{d_{ij}^o}{20 \text{ \AA}} \right)^2} \quad (7)$$

Hierbei ist d_{ij}^o der mittlere Abstand der Reste (eine standardisierte Version des Scores, so wird also ebenfalls der Z-Score³⁴ berechnet). Die Z-Werte für alle Proteinepaare werden berechnet. Dabei wird eine Z-Score-Matrix erstellt, wobei

³³ [Merkel, 2009]

³⁴ Der Z-Score ist ein statistisches Maß, dass den Abstand (gemessen in Standardabweichungen) eines Datenpunktes aus dem Mittelwert eines Datensatzes quantifiziert.

ähnliche Proteine nah beieinander und unterschiedliche weit auseinander liegen. Obwohl die Z-Werte einige Aspekte der strukturellen Ähnlichkeit quantifiziert, müssen Details dieser Metrik sorgfältig geprüft werden. Da beispielsweise die Z-Werte aus paarweisen Alignments erzeugt werden, können verschiedene Residues für jedes Paar verwendet werden. So stehen die Werte der Z-Score-Matrix für verschiedene Gesichtspunkte der strukturellen Ähnlichkeit und sind nicht direkt miteinander vergleichbar.³⁵

3.3.2 PyMOL

PyMOL ist ein Programm zur Visualisierung von Molekülstrukturen. Die Protein-daten können beispielsweise von einer Datenbank heruntergeladen und mit Hilfe von PyMOL geöffnet werden. PyMOL bietet eine Vielzahl von Möglichkeiten, die biomolekularen Strukturen zu betrachten und bestimmte Teile und Abschnitte eines Proteins detailliert darzustellen. Mit PyMOL lassen sich in begrenztem Maße Veränderungen an Proteinen, zum Beispiel durch Punktmutationen, und deren Auswirkungen auf die Struktur beobachten. Die Anwendung ist daher ein nützliches Werkzeug in der Biologie, Bioinformatik und Pharmazie.³⁶

3.3.3 MEPAL

MEPAL (multiple energy profile alignment) ist ein Algorithmus, der, wie andere klassische multiple Alignmentverfahren, auf eine Heuristik zur Approximation des optimalen multiplen Alignments zurückgreift. Eine Heuristik wird verwendet, weil ohne sie der Zeitaufwand und Speicherbedarf zu groß wäre.

³⁵ [Guerra, 2003]

³⁶ [Rau et al., 2010]

Die Berechnung des optimalen Alignments zweier Energieprofile entspricht der Berechnung einer zweidimensionalen Optimierungsmatrix, dabei ist die Größe der Matrix abhängig von der Länge der Energieprofile. Weiterhin gilt, dass die Dimension des Optimierungsproblems mit der Anzahl der Eingaben steigt. Da eine Lösung nicht in annehmbarer Zeit berechnet werden kann, ist das Problem NP-vollständig. Um diese Problematik lösen zu können, benötigt man Heuristiken. Der MEPAL-Algorithmus beruht auf der Basis einer klassischen Lösungsheuristik, wie man es bei CLUSTAL kennt. Er bearbeitet vorab zwei Schritte. Zunächst wird eine Distanzmatrix für alle eingegebenen Daten ermittelt, woraus dann eine Ableitung eines Guide Trees hervorgeht. Damit kann nun die Reihenfolge der zu alignenden Profile ermittelt und das optimale Alignment gefunden werden.

Zur Berechnung wird ein Alignment-Algorithmus verwendet, der in [Heinke, 2011] nachzulesen ist. Ausgehend davon wird das Multiple-Energieprofil-Alignment berechnet und einige Parameter dem multiplen Alignment angepasst.

$$D := \text{dScore}(S_a; S_b) = -\log_{10} \left(\frac{(x_r - x_{rand})}{x_{opt} - x_{rand}} \right) \quad (7)$$

Bei der Berechnung der Distanzmatrix wird eine Menge S , für die Energieprofile $\{s_1, s_2, \dots, s_n\}$, angenommen. Die Matrix ist mit $D_{n \times n}$ definiert. Weiterhin ist die Algorithmusfunktion zweier Energieprofile gegeben als $f(s_i; s_j)$ und der daraus sich ergebende Score mit $\text{dScore}_f(s_i; s_j)$. ϵ_{open} und ϵ_{extend} werden als konstant angenommen. Für alle Paarungen $(s_i; s_j \in S)$ ist $D_{i,j} = \text{dScore}_f(s_i; s_j)$. Dann ist D die Distanzmatrix von S . Weil die Matrix symmetrisch ist, besitzt die Hauptdiagonale nur Nullbeträge.

Die Ermittlung des Guide Trees geschieht mittels eines binären Baumes. Dabei ist ein binärer Baum eine spezielle Form eines Graphen, bei dem jeder Knoten maximal zwei Kindknoten besitzt. Ein binärer Baum ist entweder leer oder besteht aus Knoten, an denen wiederum Bäume (Unterbäume) hängen. Weiterhin ist ein Cluster c definiert durch einen Knoten und seine zwei Kanten.

Aber auch ein Blatt ist ein Cluster ohne wegführende Kanten. Somit ergibt sich, dass der Guide Tree selbst auch ein Cluster darstellt. Daraus geht hervor, dass ein Guide Tree aus einer Menge von Cluster C besteht. Die Cluster stehen dabei mit der Relation „Kanten“ ganz oder teilweise in Beziehung. Die binäre Baumstruktur wird somit aus den Blättern und Cluster und deren Relationen zueinander aufgebaut. Dabei werden die Cluster verschiedenartig zusammengefasst. Zunächst wird der kleinste Score aller möglichen Paare von Cluster in C gesucht und der kleinste Wert dieser drei Distanzen wird ermittelt.

Die Cluster werden in einer Clusterliste wie folgt zusammengefasst:

- zwei ungeclusterte Energieprofile werden zu einem Cluster
- oder ein Energieprofil wird zu einem bestehenden Cluster hinzugefügt
- oder zwei bestehende Cluster werden zu Tochter-Cluster, über einen Vaterknoten verbunden und so zu einem Cluster vereinigt

Diese Liste beinhaltet die Veränderungen im Baum pro Iteration.

Ein Guide Tree zeigt, im Gegensatz zu einem phylogenetischen Baum, keine evolutionären Beziehungen, sondern gibt in diesem Fall die Distanzen zwischen den Energieprofilen wieder.

Um das Alignment zu berechnen, wird die Clusterliste herangezogen. Die in der Liste abgespeicherten Veränderungen werden nun progressiv durchlaufen. Die zwei Profile, die in einem Cluster liegen, werden align und abgespeichert. Die Lückenstrafen und Lückeneröffnungsstrafen sind äquivalent zu denen, die für die Berechnung der Matrix genommen wurden. Befindet sich im nächsten Schritt ein weiterer Cluster im Baum, so werden auch diese zwei Profile align und gespeichert. Findet jedoch im nächsten Zeitschritt eine Zusammenfügung von einem Profil mit einem Cluster statt, so wird der Centroid des Clusters ermittelt. Der Centroid ist das Profil, das den geringsten Abstand zu den anderen Energieprofilen besitzt. Dieses Centroid wird dann im darauffolgenden

Schritt mit dem ungeclusterten Energieprofil alignt. Lücken, die in den Centroiden eingeführt werden, projizieren sich in allen Profilen im Cluster. Nach dem CLUSTAL-Prinzip bleiben diese Lücken fortwährend bestehen. Diese Vorgehensweise gilt auch für zwei Cluster (Erzeugung des Centroid jedes Clusters; paarweises Alignment). Schließlich kann so das globale Energieprofil-Alignment ermittelt werden.

Möchte man zudem noch das Consensus-Profil ermitteln, wird eine Art Centroid-Verfahren zum Einsatz gebracht. Es muss der Energiewert jedes Alignments gesucht werden, der den besten mittleren Score aufweist, gegenüber den anderen Alignments. Durch das Zusammenführen aller Centroid-Energien entsteht das Consensus-Profil.³⁷

37 [Heinke, 2011]

4 Strukturanalyse

Die für die Strukturanalyse benötigten Daten wurden der PDB entnommen. Hierfür wurden zwei Datensätze mit Hilfe dieser Proteindatenbank angefertigt. Es entstand ein Datensatz mit dem globulären Protein Lysozym und ein Datensatz mit dem Membranprotein Bacteriorhodopsin. Zu jedem Wildtyp wurden zehn Mutanten mit einer Punktmutation mit Hilfe der Literaturangaben in der PDB und Pubmed erstellt. Mit der Literatur sollten die Auswirkungen der Punktmutationen auf das Protein herausgefunden werden. Danach wurde der Abstand der einzelnen Aminosäuren des Wildtyps mit seiner Mutanten direkt verglichen. Mittels der RMSD-Werte (root mean square deviation) der einzelnen Aminosäuren wurden so für jede Mutante ein Diagramm erzeugt, in denen man die Auswirkungen der Mutationen auf die Struktur erkennen kann. Zur Veranschaulichung werden zudem die Sekundärstrukturen des Wildtyps und der Mutante unter der Abstandskurve abgebildet.

Weiterhin wird die Wahrscheinlichkeit der Mutationen auf Codon-Ebene untersucht, d.h. man möchte wissen, welche Mutationen in der Natur theoretisch häufiger auftreten als andere. Die Ergebnisse siehe Anhang (Tabelle 8 und Tabelle 9) zu finden.

Außerdem wird ein Multistrukturalignment durchgeführt, um die Strukturähnlichkeiten aller Mutanten gegenüber dem Wildtyp untersuchen zu können. Weiterhin wird aus den Q-Scores der Mutanten eine Distanzmatrix erstellt, um mit dieser einen Distanzbaum darstellen zu können.

4.1 Literaturrecherche bezüglich der Mutationsauswirkungen

Dem Wildtyp Bacteriorhodopsin (PDB-ID: 2NTU) werden zehn Mutationen, die eine Punktmutation aufweisen, gegenübergestellt und die Auswirkungen auf die Funktion, Stabilität und Struktur untersucht. Die Literaturrecherche hat folgende Ergebnisse hervorgebracht:

Tabelle 4: Auswirkungen der Punktmutationen in Hinblick auf Stabilität, Struktur und Funktion (Bacteriorhodopsin) ^{38 39 40 41 42 43 44}

PDB-ID	Mutation	Stabilitätsänderung	Strukturänderung	Funktionsverlust
2I1X	D96A	--	--	--
2I21	T46V	-	-	-
1TN0	A51P	---	---	--
3HAP	L111A	-	---	-
1PXS	M56A	+	---	keine
1Q5I	P186A	-	---	-
1PXR	P50A	-	---	-
1P8I	F219L	+	-	keine
2WJL	E194D	--	--	--
3HAQ	I148A	--	---	--

Dabei stehen die - Zeichen für eine Verschlechterung der Eigenschaften und ein + Zeichen für eine verbesserte Eigenschaft. Dabei wurde die Bewertung subjektiv vorgenommen.

38 [Lanyi et al., 2006]

39 [Yohannan, 2004]

40 [Joh, 2009]

41 [Faham, 2004]

42 [YohannanII, 2004]

43 [Schobert, 2003]

44 [Garczarek, 2005]

Äquivalent wurde diese Recherche mit dem globulären Protein Lysozym vorgenommen (siehe Tabelle 5), wobei hier der Wildtyp die PDB-ID 3LZM trägt.

Tabelle 5: Auswirkungen der Punktmutationen in Hinblick auf Stabilität, Struktur und Funktion (Lysozym) ^{45 46 47}

PDB-ID	Mutation	Stabilitätsänderung	Strukturänderung	Funktionsverlust
3C7Y	R96A	-	-	-
3C80	R96Y	- - -	- -	- -
3C81	K85A	+ - (je nach pH-Wert)	-	-
3C83	D89A	-	-	-
3CDR	R96Q	-	-	-
3C8R	R96G	-	-	-
3C8S	R96E	- (je höher der pH-Wert desto destabilisierender)	-	-
3CDQ	R96S	-	- -	- -
3C7W	R96K	-	-	- -
3F8V	R96H	- -	- -	- -

4.2 Strukturvergleich mittels Dali-Server

Abbildung 13 stellt ein Multistrukturalignment für Bacteriorhodopsin und die dazugehörigen zehn Mutanten dar und Abbildung 14 zeigt das Multistrukturalignment für Lysozym mit deren zehn Mutanten. Man erkennt die strukturellen Unterschiede der ausgewählten Proteine und deren Mutanten. Die größten Unregelmäßigkeiten befinden sich jeweils am C-Terminus, da hier die Endcoils liegen.

⁴⁵ [Collins, 2005]

⁴⁶ [Mooers, 2009]

⁴⁷ [MooersII, 2009]



Abbildung 13: Multistrukturalignment (Bacetriforhodopsin) mit Hilfe des Dali-Servers, dabei kann man aus einer Liste von strukturrähnlichen Proteinen die gewünschten zu vergleichenden Strukturen auswählen

Weiterhin ist zu sehen, dass sich die Helices und Sheets durch die Punktmutationen entweder verkürzen oder auch verlängern.

4.3 Strukturvergleich mittels RMSD-Diagramms

Strukturelle Änderungen lassen sich über Root Mean Square Displacements verfolgen. Hierzu wird jede Mutante einzeln mit dem Wildtyp verglichen, indem zunächst RMSD-Diagramme erstellt werden. Mit Hilfe des RMSD Calculators⁴⁸ werden die einzelnen Abstände der Aminosäuren vom Wildtyp und der Mutante in Ångström (Å) berechnet (siehe Abbildung 15) und können danach in einem Diagramm veranschaulicht werden. In Abbildung 16 und 17 sieht man je ein Beispieldiagramm für die untersuchten Proteine, das zusätzlich einen direkten Strukturvergleich beinhaltet. Diagramme für die anderen Mutanten sind im Anhang zu finden.



Abbildung 15: RMSD-Berechnung mit RMSD Calculator.
Vergleich von Bacteriorhodopsin und die Mutante 3HAP

48 <http://www.igs.cnrs-mrs.fr/Caspr2/RMSDcalc.cgi?step=0>

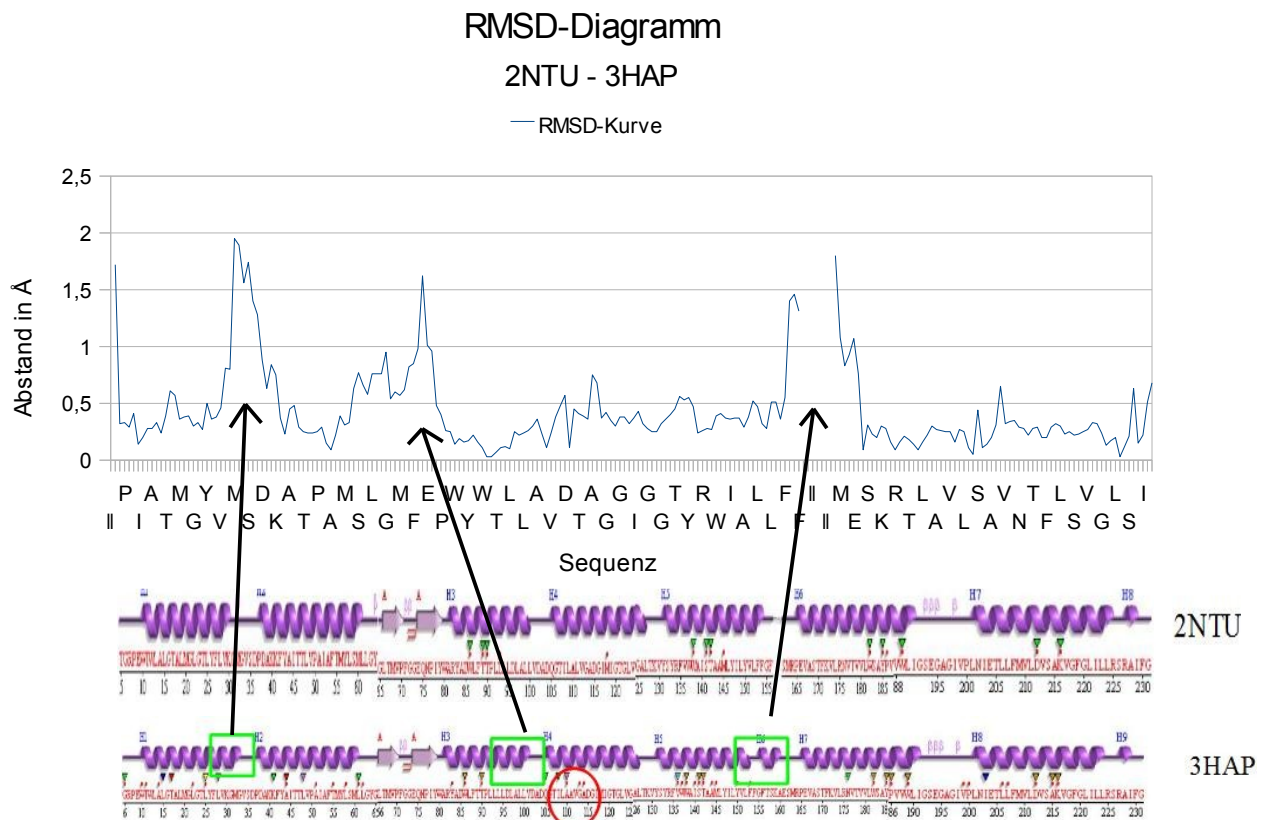


Abbildung 16: Bacteriorhodopsin-RMSD-Diagramm: Wildtyp(2NTU)-Mutante(3HAP)

Die grünen Quadrate zeigen die Stellen, an denen eine strukturelle Veränderung stattfindet und der rote Kreis markiert die Mutationsstelle.

Im RMSD-Diagramm ist der Vergleich des Wildtyps von Bacteriorhodopsin und der Mutante 3HAP zu sehen. Die Peaks im Diagramm stellen die Veränderungen der Struktur in der Mutante dar. Der Ausschlag am Anfang kann vernachlässigt werden, weil es sich hier um die sich am Anfang befindlichen Coils handelt, die das Protein nicht stark beeinflussen. Die Lücke in der Kurve zeigt einen nicht aufgelösten Bereich des Wildtyps an, die jedoch in der Mutante geschlossen ist.

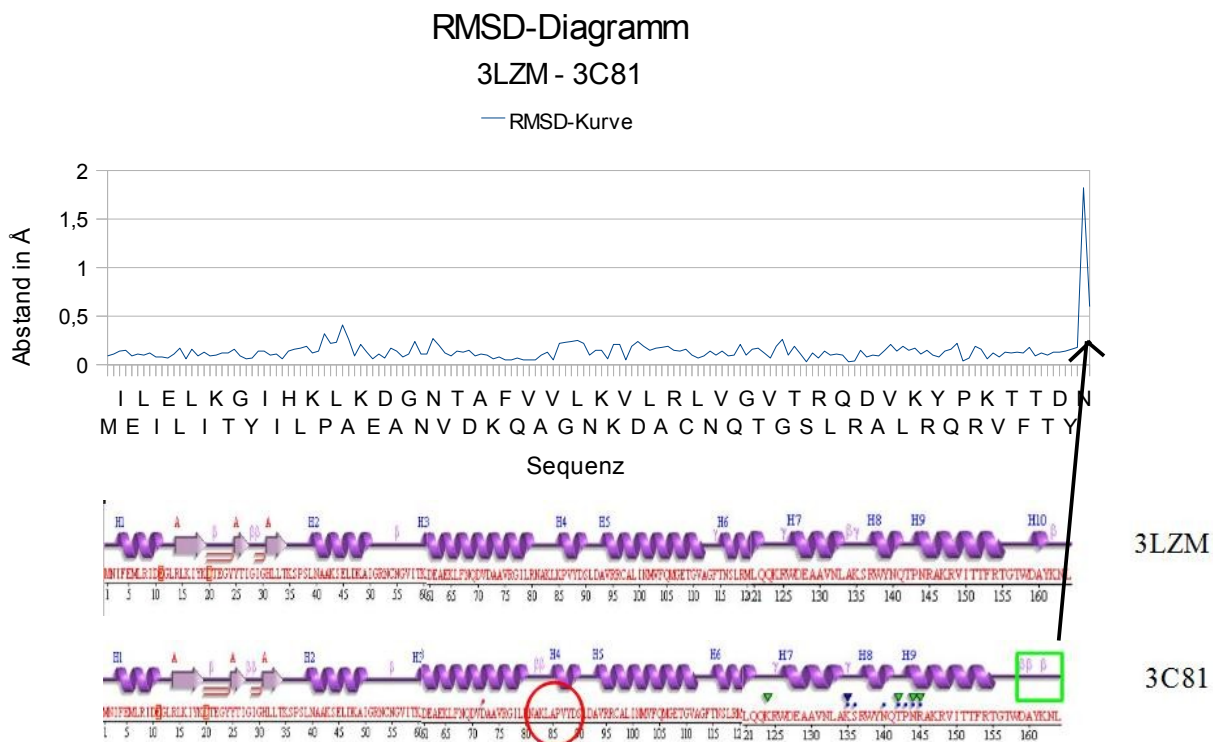


Abbildung 17: Lysozym-RMSD-Diagramm: Wildtyp(3LZM)-Mutante(3C81)

Das grüne Quadrat zeigt die Stelle, an der eine strukturelle Veränderung stattfindet und der rote Kreis markiert die Mutationsstelle.

Im RMSD-Diagramm ist der Vergleich des Wildtyps von Lysozym und der Mutante 3C81 zu sehen. Es ist nur eine strukturelle Veränderung zu erkennen, die jedoch außer Acht gelassen werden kann, da sie sich nicht auf das Protein auswirkt. Insgesamt ist zu sagen, dass sich die Mutante durch die Punktmutation nur minimal vom Wildtyp unterscheidet.

4.4 Strukturvergleich mittels Distanzbaum

Um die Struktur mittels eines Distanzbaumes zu untersuchen, ist dazu eine Distanzmatrix nötig. Diese Distanzmatrizen wurden mit Hilfe des Dali-Servers ermittelt (siehe Tabelle 6 und 7). Das durchgeführte Multistrukturalignment berechnete Q-Werte, die die Strukturähnlichkeit der Mutanten und des Wildtyps angibt. Da man den Abstand zwischen den Proteinen benötigte, musste man 1- Q-Wert rechnen. Mit dieser entstandenen Abstandsmatrix wurde mit dem Programm Mobyli@Pasteur der Distanzbaum erstellt. Dieses Programm benutzt dafür die Neighbour Joining Methode.

Tabelle 6: Q-Score-Tabelle von Bacteriorhodopsin berechnet mit dem Dali-Server

Struktur		1	2	3	4	5	6	7	8	9	10	11
1	2NTU		0,978	0,98	0,923	0,936	0,929	0,926	0,932	0,984	0,975	0,934
2	2I1X	0,978		0,988	0,921	0,933	0,927	0,926	0,931	0,983	0,967	0,931
3	2I21	0,98	0,988		0,924	0,935	0,93	0,928	0,934	0,985	0,968	0,933
4	1TN0	0,923	0,921	0,924		0,942	0,942	0,943	0,944	0,925	0,927	0,942
5	3HAP	0,936	0,933	0,935	0,942		0,945	0,945	0,949	0,937	0,939	0,955
6	1PXS	0,929	0,927	0,93	0,942	0,945		0,944	0,948	0,932	0,931	0,945
7	1Q5I	0,926	0,926	0,928	0,943	0,945	0,944		0,948	0,927	0,929	0,945
8	1PXR	0,932	0,931	0,934	0,944	0,949	0,948	0,948		0,934	0,934	0,949
9	1P8I	0,984	0,983	0,985	0,925	0,937	0,932	0,927	0,934		0,972	0,935
10	2WJL	0,975	0,967	0,968	0,927	0,939	0,931	0,929	0,934	0,972		0,938
11	3HAQ	0,934	0,931	0,933	0,942	0,955	0,945	0,945	0,949	0,935	0,938	

Tabelle 7: Q-Score-Tabelle von Lysozym berechnet mit dem Dali-Server

Struktur		1	2	3	4	5	6	7	8	9	10	11
1	3LZM		0,972	0,977	0,974	0,974	0,973	0,971	0,972	0,972	0,973	0,963
2	3C7Y	0,972		0,979	0,973	0,973	0,973	0,97	0,973	0,972	0,972	0,964
3	3C80	0,977	0,979		0,978	0,979	0,981	0,98	0,981	0,981	0,979	0,978
4	3C81	0,974	0,973	0,978		0,975	0,974	0,971	0,973	0,973	0,974	0,965
5	3C83	0,974	0,973	0,979	0,975		0,974	0,971	0,974	0,974	0,974	0,965
6	3CDR	0,973	0,973	0,981	0,974	0,974		0,973	0,974	0,974	0,975	0,965
7	3C8R	0,971	0,970	0,98	0,971	0,971	0,973		0,972	0,972	0,972	0,959
8	3C8S	0,972	0,973	0,981	0,973	0,974	0,974	0,972		0,974	0,974	0,967
9	3CDQ	0,972	0,972	0,981	0,973	0,974	0,974	0,972	0,974		0,974	0,967
10	3C7W	0,973	0,972	0,979	0,974	0,974	0,975	0,972	0,974	0,974		0,964
11	3F8V	0,963	0,964	0,978	0,965	0,965	0,965	0,959	0,967	0,967	0,964	

In den Distanzbäumen werden die strukturellen Unterschiede aller Mutanten und dessen Wildtyp dargestellt. Dabei zeigt Abbildung 18 den Baum des Proteins Bacteriorhodopsin und Abbildung 19 den Distanzbaum des Proteins Lysozym. Außerdem sind die Tabellen aus der Literaturrecherche aus Kapitel 4.1 darübergerlegt.

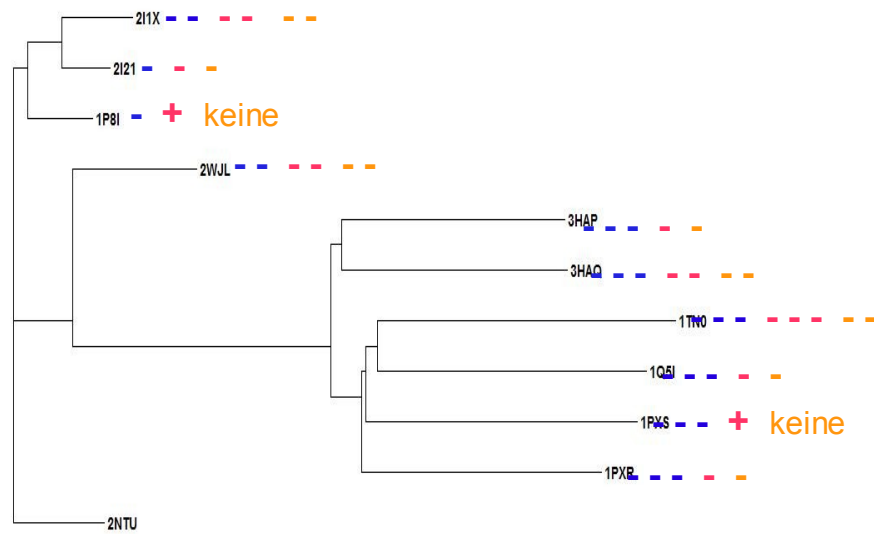


Abbildung 18: Distanzbaum von Bacteriorhodopsin mit der Ausgabe von Mobyle@Pasteur

Der Distanzbaum wurde zusätzlich mit den Ergebnissen der Literaturrecherche gelabelt, wobei blau die Strukturveränderung, rosa die Stabilitätsänderung und orange den Funktionsverlust symbolisiert.

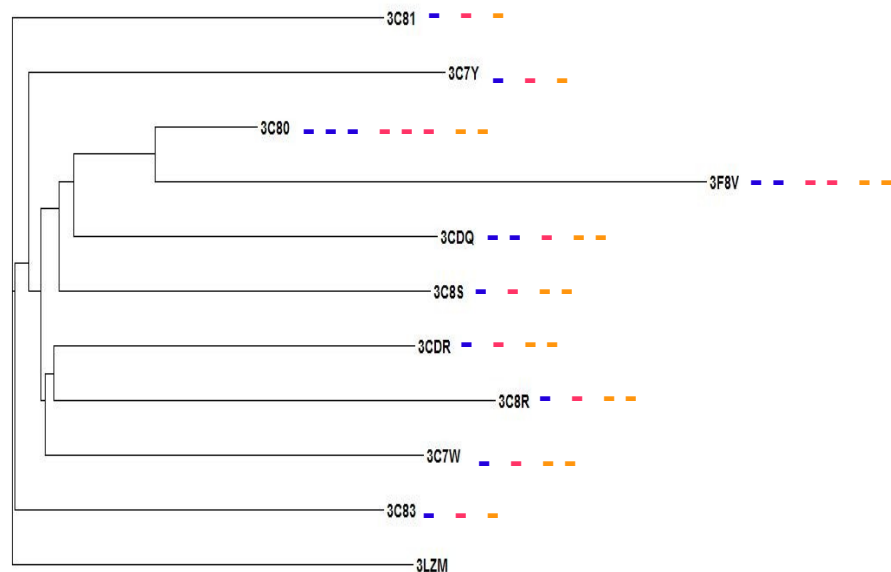


Abbildung 19: Distanzbaum von Lysozym mit der Ausgabe von Mobyle@Pasteur

Der Distanzbaum wurde zusätzlich mit den Ergebnissen der Literaturrecherche gelabelt, wobei blau die Strukturveränderung, rosa die Stabilitätsänderung und orange den Funktionsverlust symbolisiert.

5 Energieprofilanalyse

5.1 Analyse mittels Energieprofilalignment

Die paarweisen Strukturalignments und Multistrukturalignments geben Aufschluss über Bereiche mit ähnlichen räumlichen Koordinaten. Starke Abweichungen von wichtigen Strukturelementen können die Proteinfunktion und ihre Faltung beeinflussen. Energieprofile verknüpfen sowohl Rauminformationen als auch die Charakteristika der Aminosäuren. Somit stellt ein Energieprofil-Alignment die Abschnitte dar, deren Aminosäurereste ähnliche signifikante Wechselwirkungen besitzen.

Die benötigten Energiewerte wurden mit Hilfe des Hochschulservers eProS ermittelt. Dabei mussten die Werte für jedes Protein einzeln berechnet werden. Alle verwendeten Proteine sind in der PDB mit der Sequenz und ihren Koordinaten gespeichert.

Ein Energieprofilalignment zeigt die Energiedifferenzen der einzelnen Aminosäuren von zwei Proteinen. Zunächst müssen jedoch die Energieprofile aller ausgewählten Proteine berechnet werden. Zwei Beispielennergieprofile, die später verglichen werden, sind in Abbildung 20 und 21 zu sehen. Dabei handelt es sich um den Wildtyp Bacteriorhodopsin und eine Mutante. Äquivalent dazu werden in Abbildung 23 und 24 der Wildtyp von Lysozym und einer seiner Mutante dargestellt. Im Anschluss daran werden diese Paare mittels dem Energieprofilalignment miteinander verglichen (siehe Abbildung 22 und 25). Alle anderen Energieprofilealignments befinden sich im Anhang.

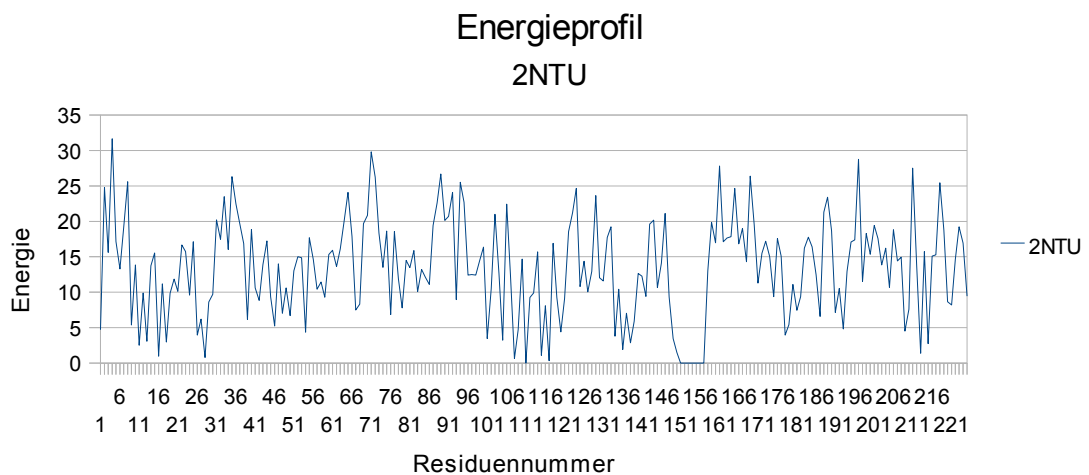


Abbildung 20: Energieprofil des Wildtyps Bacteriorhodopsin

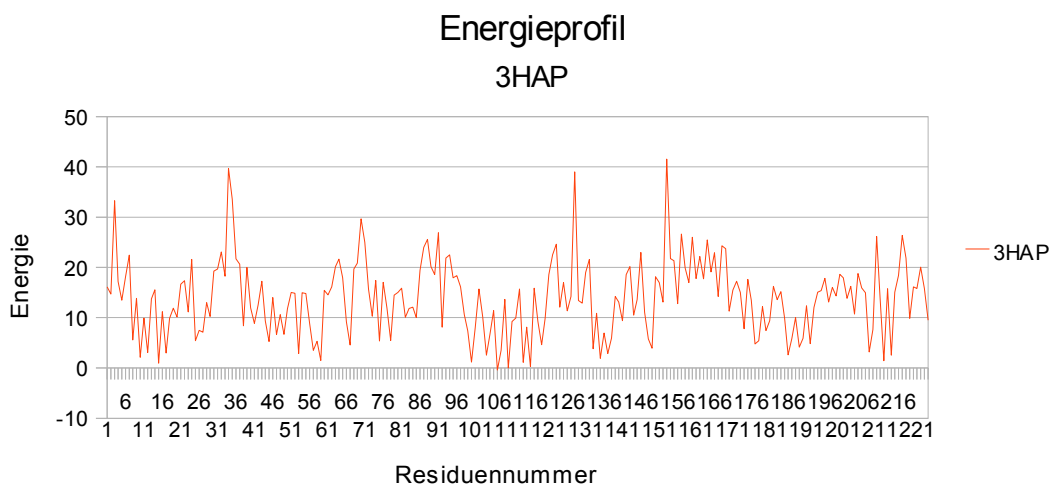


Abbildung 21: Energieprofil der BR-Mutante 3HAP

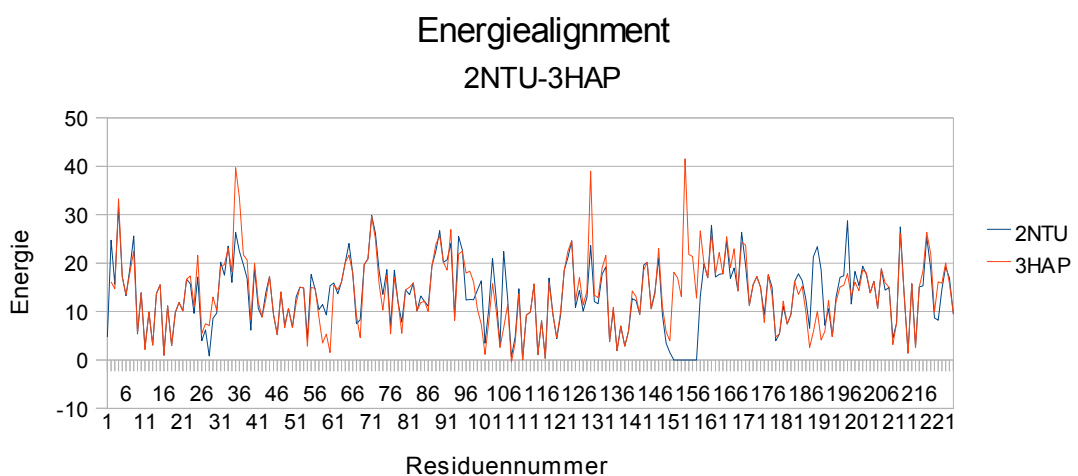


Abbildung 22: Energieprofilalignment von 2NTU und 3HAP

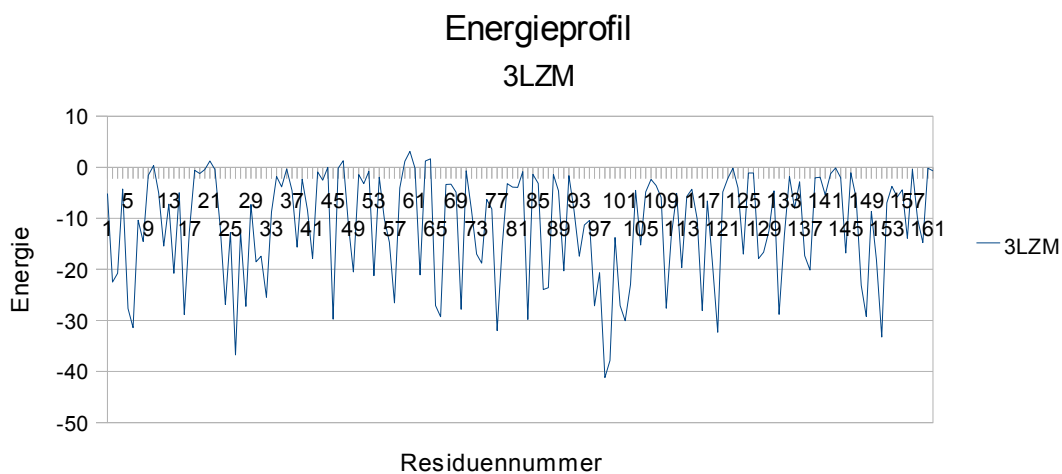


Abbildung 23: Energieprofil des Wildtyps Lysozym

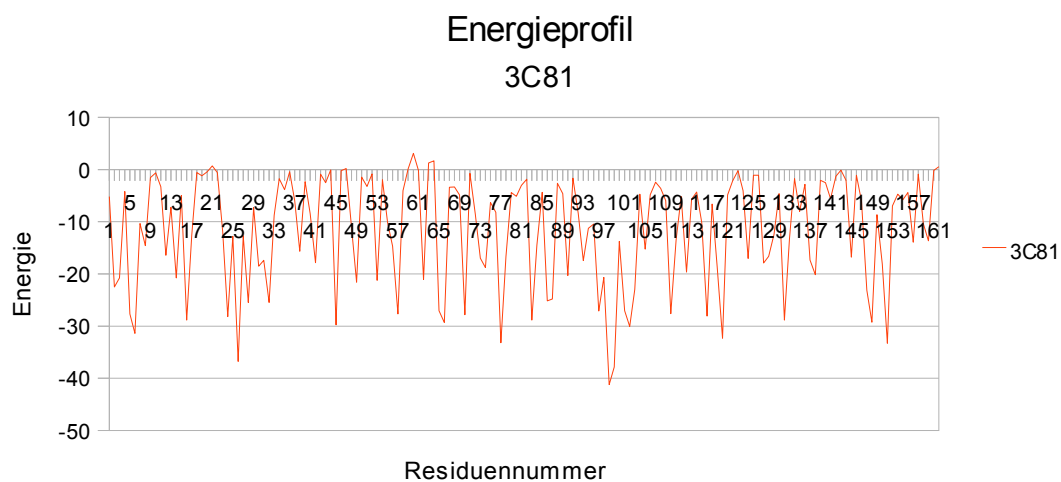


Abbildung 24: Energieprofil der Mutante 3C81

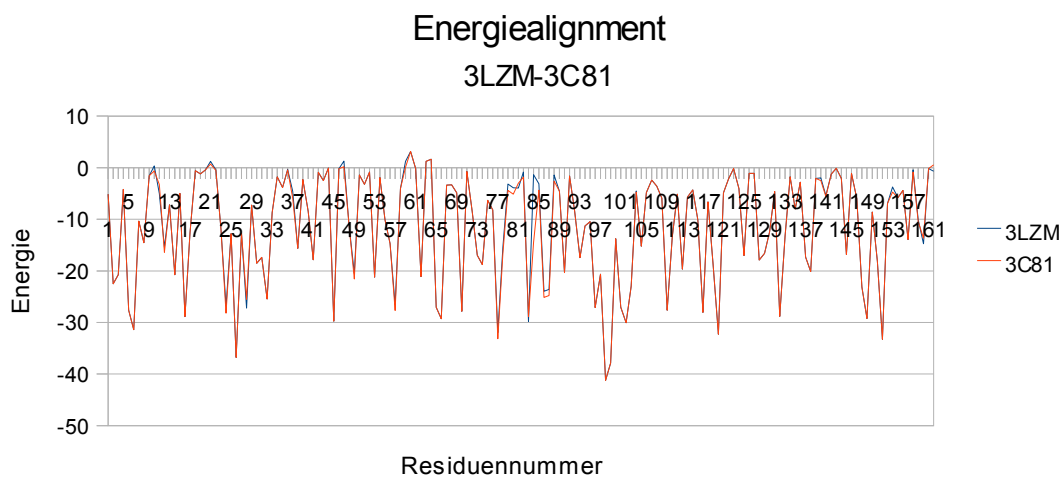


Abbildung 25: Energieprofilalignment von 3LZM und 3C81

Das Tool eMut des eProS-Servers bietet die Möglichkeit, sich die durch die Punktmutation verursachten Veränderungen im Energieprofil anzeigen zu lassen. Abbildung 26 und 27 zeigen die Mutationsenergieprofile der oben ausgewählten Beispiele. Alle weiteren Mutationsenergieprofile sind im Anhang zu finden.

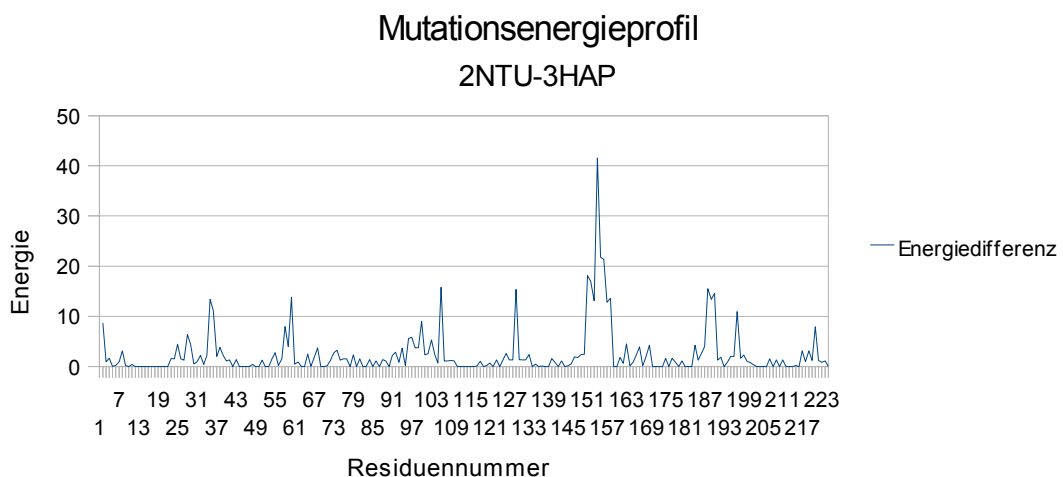


Abbildung 26: Mutationsenergieprofil von 2NTU und 3HAP

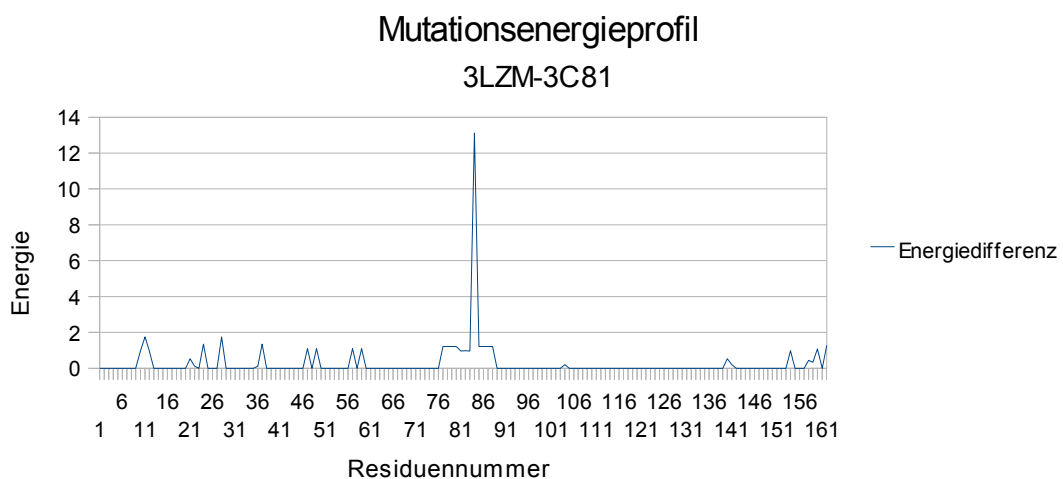


Abbildung 27: Mutationsenergieprofil von 3LZM und 3C81

5.2 Analyse mittels MEPAL

MEPAL ist ein Programm, welches mir freundlicherweise von Florian Heinke zur Verfügung gestellt wurde. Es steht für multiple energy profile alignment und vergleicht somit alle Energieprofile eines ganzen Proteinsatzes.

5.2.1 Multienergiealignments

In Abbildung 28 ist das Multienergiealignment von Bacteriorhodopsin zu sehen und äquivalent dazu in Abbildung 29 das multiple Alignment von Lysozym. Die obersten Reihen stellen die direkten Energieprofilalignments dar. Dabei ist jede Residue durch den One-letter-code gelabelt. Weiterhin repräsentiert die Färbung die Energie der Residue, blau steht für eine niedrige und rot für eine hohe Energie. Die mittlere Zeile visualisiert das Consensus-Energieprofil und die unterste Spalte zeigt die energetische Konservierung für jede Position im Alignment an. Es ist zu sagen, dass einige BRs Lücken (fehlende Resiudes und Koordinaten) aufweisen. Diese konnten wahrscheinlich im X-ray⁴⁷ nicht eindeutig aufgelöst werden. Die Berechnung der Energieprofile von Membranproteinen kann nur erfolgen, wenn membranständige Bereiche bekannt sind (siehe [Brumm, 2010]). Hierfür wird die Annotation des Proteins auf der PDBTM (Protein Data Bank of Transmembrane Proteins) abgerufen und zur Berechnung herangezogen. Da jedoch diese Annotation für meine ausgewählten BRs nicht eindeutig ist (und somit die Berechnung mit deren Hilfe nicht möglich ist), wurde dies mittels des TMDet-Servers⁴⁸ (Detection of transmembrane regions), der die membranständigen Bereiche der Proteinstrukturen sehr exakt vorhersagen kann, berechnet.

47 X-ray, auch Röntgenstrahlung genannt, bezeichnet elektromagnetische Wellen mit Photonenenergien zwischen 100 eV und einigen MeV.

48 Der TMDet-Web-Server dient zur Erkennung von transmembranen Regionen von Proteinen, mit Hilfe ihrer 3D-Koordinaten.

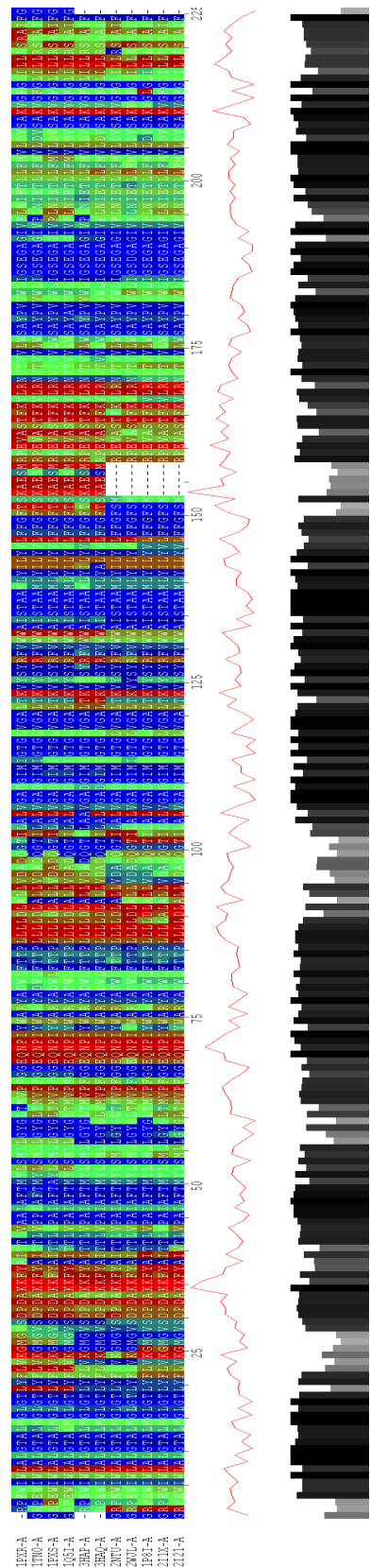


Abbildung 28: Multienergiealignment von Bacteriorhodopsin

Die obersten Reihen stellen die direkten Energieprofilalignments dar, wobei blau eine niedrige und rot eine hohe Energie symbolisiert. Die rote in der Mitte liegende Kurve visualisiert das Consensus-Energieprofil und die unterste Spalte zeigt die energetische Konservierung für jede Position im Alignment an.

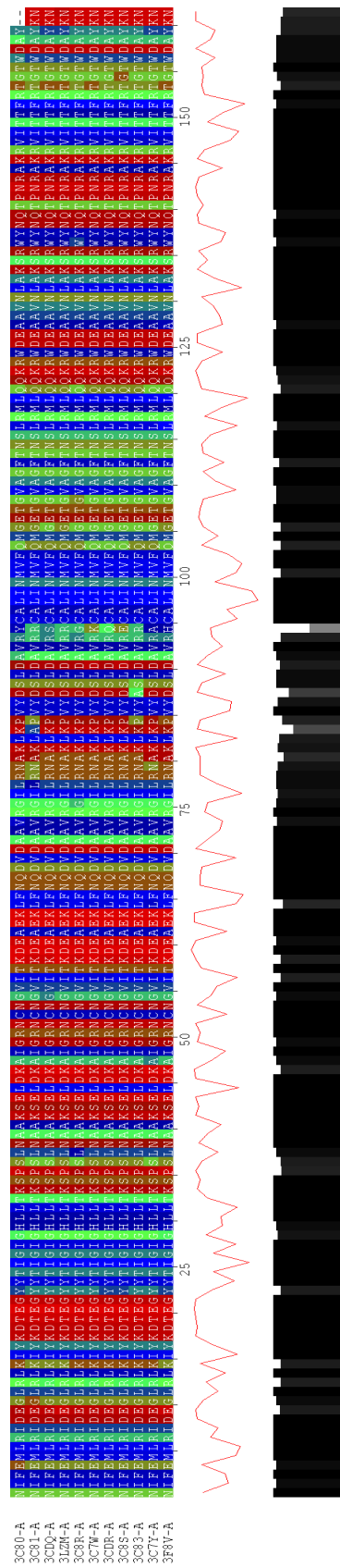


Abbildung 29: Multienergyalignment von Lysozym

Die obersten Reihen stellen die direkten Energieprofilalignments dar, wobei blau eine niedrige und rot eine hohe Energie symbolisiert. Die rote in der Mitte liegende Kurve visualisiert das Consensus-Energieprofil und die unterste Spalte zeigt die energetische Konservierung für jede Position im Alignment an.

Bei den Alignments wurden die Lückeneröffnungsstrafe bei -40 und die Lückenerweiterungsstrafe bei -7 festgelegt.

5.2.2 Energiedistanzbäume

Die aus dem Multienergieprofilalignment resultierenden Energiedistanzbäume sind in Grafiken 30 und 31 abgebildet. Aus den Bäumen kann man nun die Beziehungen zwischen den einzelnen Energieprofilen ablesen.

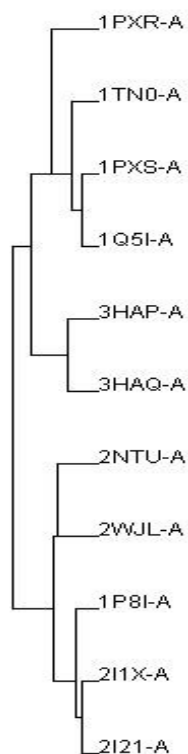


Abbildung 30: Energieprofilistanzbaum von Bacteriorhodopsin

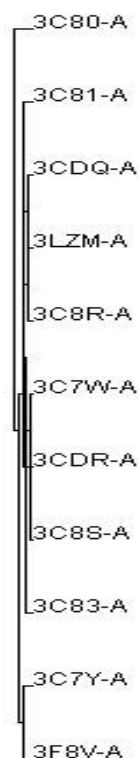


Abbildung 31: Energieprofilistanzbaum von Lysozym

6 Diskussion

6.1 Vergleich der RMSD- und Mutationsenergieprofilendiagramme

Vergleicht man das RMSD-Diagramm mit den dazugehörigen Mutationsenergieprofilendiagramm aus Kapitel 4.3 und 5.1, so erkennt man, dass sich die Diagramme in der Grafik 32 sehr ähneln, was eine Beziehung zwischen der Strukturänderung (im RMSD-Diagramm zu sehen) und der Energiedifferenz erraten lässt. Im Gegensatz dazu stimmen die Kurven des Beispiel-Lysozym-Paares nicht überein.

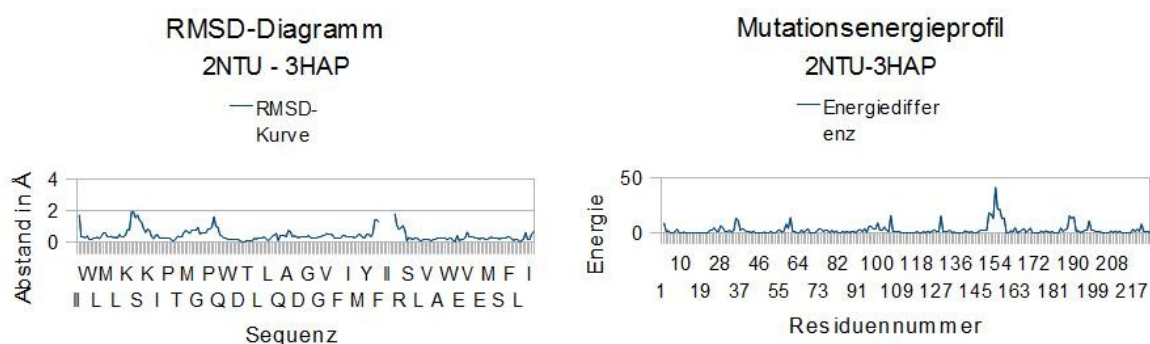


Abbildung 32: Vergleich RMSD-Diagramm und Mutationsprofil von 2NTU-3HAP

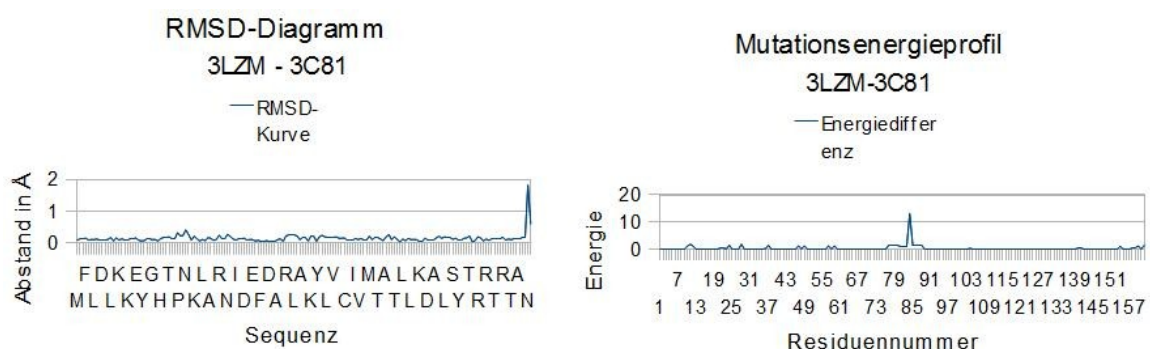


Abbildung 33: Vergleich RMSD-Diagramm und Mutationsprofil von 3LZM-3C81

Nun stellt sich die Frage, ob man einen Zusammenhang zwischen der Strukturänderung und der Energiedifferenz findet. Dafür wird der erstellte Distanzbaum und der Guide Tree der Energieprofile miteinander verglichen.

6.2 Vergleich von Distanzbaum und Energieprofilidistanzbaum

In Abbildung 34 sind die beiden Distanzbäume von Bacteriorhodopsin gegenübergestellt. Dabei basiert der linke Baum auf den RMSD-Werten und der rechte auf den Energiedifferenzen. Man erkennt, dass die Gruppierungen in beiden Grafiken vorhanden sind. Dies lässt den Schluss zu, dass jede Punktmutation sich im gleichen Maße auf die Struktur, als auch auf die Energie und somit auf die Stabilität des Proteins auswirkt.

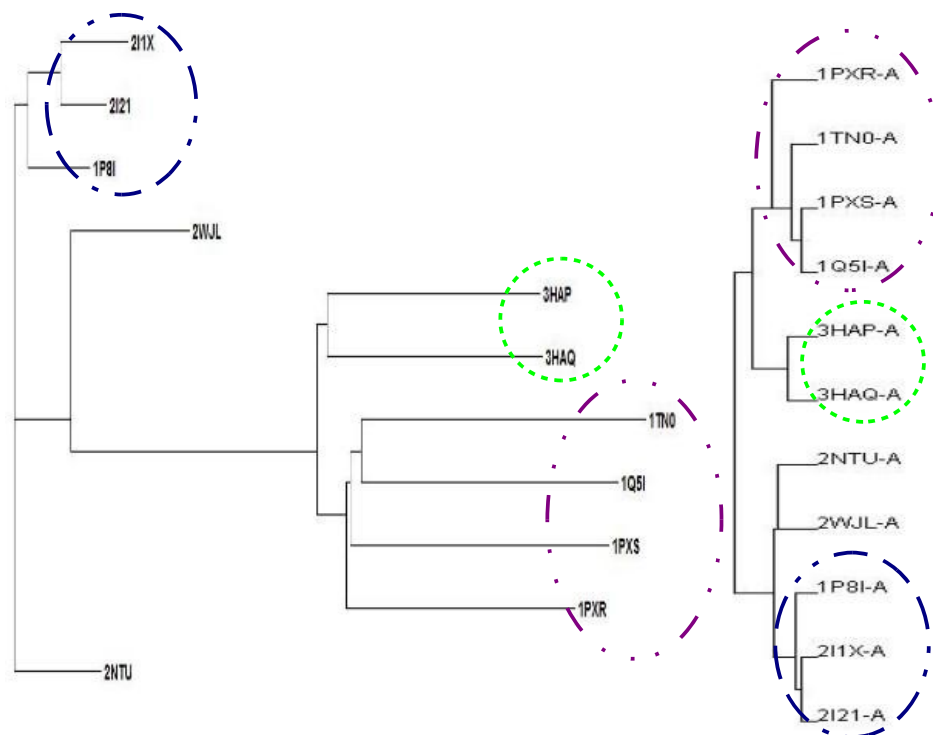


Abbildung 34: Vergleich von Distanzbaum und Energiedistanzbaum (Bacteriorhodopsin)

Links ist der Distanzbaum basierend auf den RMSD-Werten und rechts sieht man den Energiedistanzbaum, der auf den Energiedifferenzen beruht. Die verschiedenfarbige Kreise markieren die gleichen Gruppierungen.

Weiterhin geht aus beiden Abbildungen hervor, dass Mutanten, die weit entfernt vom Wildtyp 2NTU im Baum liegen, sich strukturell und energetisch stärker unterscheiden, als jene die nah am Wildtyp sind.

Betrachtet man eine Stelle im Protein, welche eine starke strukturelle Veränderung aufweist, erkennt man, dass die Mutante in diesem Bereich weniger

Wasserstoffbrückenbindungen besitzt. Dafür wurden die Wasserstoffbrücken des Wildtyps und der Mutante 3HAP an den Residuen 28-33 untersucht. In Abbildung 35 sind die H-Brücken des Wildtyps weiß und die der Mutante lila dargestellt. Dies stimmt mit der Literaturrecherche überein, die eine leichte Verschlechterung der Stabilität hervorbrachte.

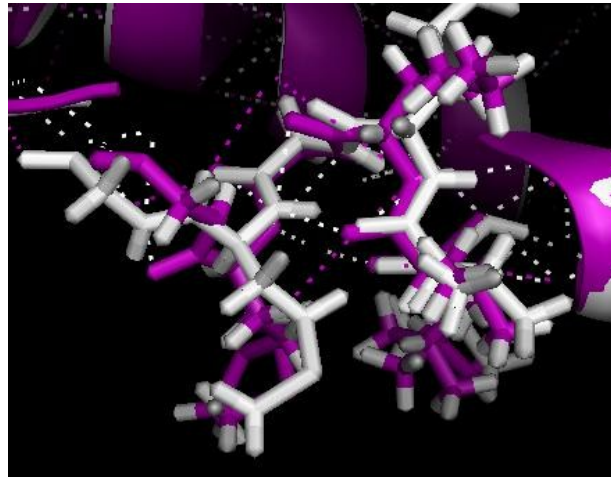


Abbildung 35: Wasserstoffbrücken des Wildtyps 2NTU (weiß) und der Mutante 3HAP (lila)

Die Untersuchung des globulären Proteins T4 Lysozym brachte ebenfalls zwei Distanzbäume hervor. Wobei der erste Guide Tree auf Grundlage der RMSD-Werte und der zweite anhand der Energiedifferenzen berechnet wurde. Abbildung 36 verdeutlicht dabei die Unterschiede der beiden Bäume. Im Gegensatz zu Bacteriorhodopsin und seiner Mutanten sind bei Lysozym keine identischen Gruppen zu finden. Somit lässt sich auch kein Zusammenhang zwischen der Strukturänderung und der Energie herstellen. Jedoch stimmen auch hier die Literaturangaben aus Kapitel 4.1 mit dem Distanzbaum (RMSD) und dem Energiedistanzbaum überein. Die Stabilität und die Struktur verändern sich nur geringfügig. Mit Hilfe von PyMOL wurden wiederum die Wasserstoffbrücken von dem Wildtyp 3LZM und der Mutante 3C81 visualisiert. Dabei befand sich der Bereich mit der größten strukturellen Veränderung am Residue 40. Darum wurde diese Stelle näher analysiert (siehe Abbildung 37). Da der Wildtyp mehr H-Brücken aufweist, ist dieser somit auch stabiler.

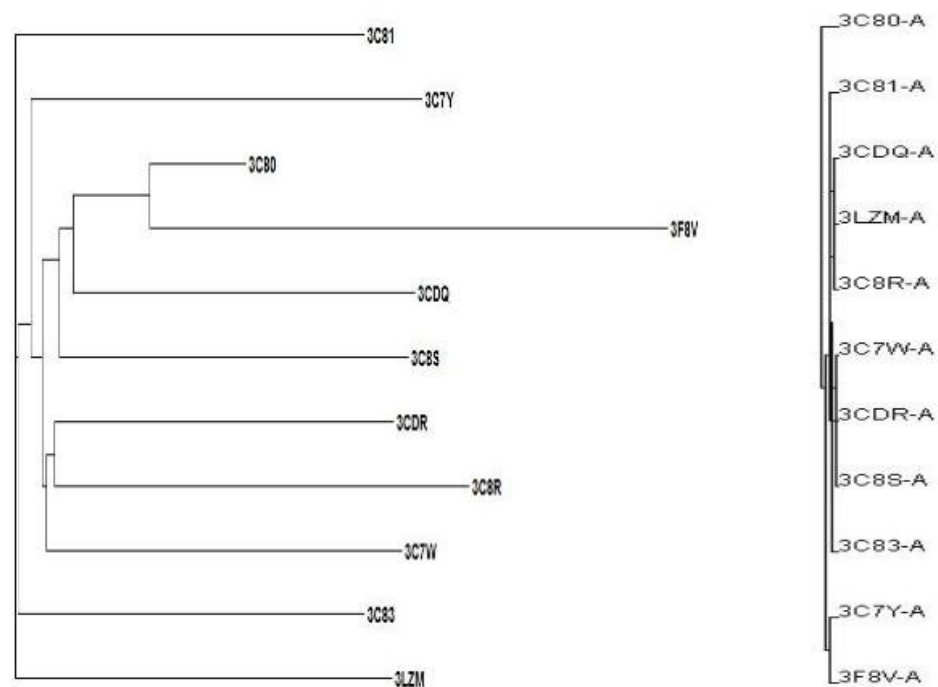


Abbildung 36: Vergleich von Distanzbaum und Energiedistanzbaum (Lysozym)

Links ist der Distanzbaum basierend auf den RMSD-Werten und rechts sieht man den Energiedistanzbaum, der auf den Energiedifferenzen beruht.

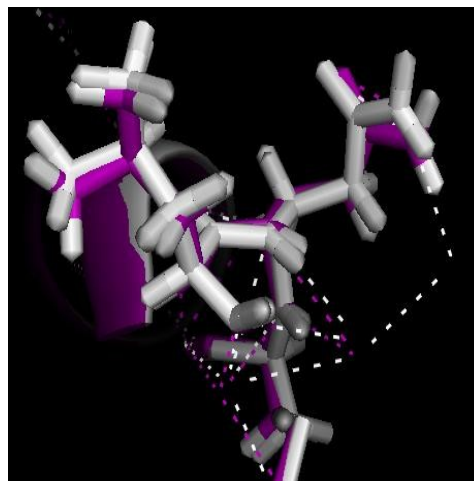


Abbildung 37: Wasserstoffbrücken des Wildtyps 3LZM (weiß) und der Mutante 3C81 (lila)

6.3 Schlussfolgerung

Es müssen noch viele Anstrengungen unternommen werden, um nachzuweisen, ob die Punktmutationen sich tatsächlich im gleichen Maße auf die Struktur und Stabilität auswirken. Weiterhin ist eine Beziehung zwischen der Strukturänderung und der Energiedifferenz nur im Membranprotein Bacteriorhodopsin nachgewiesen wurden. Demzufolge müssen weitere Untersuchungen in Bezug auf globuläre Proteine unternommen werden. Die zukünftige Forschung soll es möglich machen aus einem Energieprofil eine noch unbekannte Struktur zu realisieren.

Anhang

A.1 Mutationswahrscheinlichkeitsanalyse

Tabelle 8: Anzahl der Basenänderungen in einem Triplett, um eine Mutation im Bacteriorhodopsin auszulösen

ID	Mutation	1 Schritt nötig für Mutation	2 Schritte nötig für Mutation	3 Schritte nötig für Mutation
2I1X	D96A	GAC → GCC	GAC → GCU GAC → GCG GAC → GCA	keine
2I21	T46V	keine	ACG → GUG	ACG → GUA ACG → GUC ACG → GUU
1TN0	A51P	GCC → CCC	GCC → CCU GCC → CCG GCC → CCA	keine
3HAP	L111A	keine	CUC → GCC	CUC → GCU CUC → GCG CUC → GCA
1PXS	M56A	keine	AUG → GCG	AUG → GCC AUG → GCU AUG → GCA
1Q5I	P186A	CCC → GCC	CCC → GCU CCC → GCG CCC → GCA	keine
1PXR	P50A	CCA → GCA	CCA → GCC CCA → GCU CCA → GCG	keine
1P8I	F219L	UUC → UUG UUC → UUA	keine	keine
2WJL	E194D	GAA → GAC GAA → GAU	keine	keine
3HAQ	I148A	keine	AUC → GCC	AUC → GCU AUC → GCG AUC → GCA

Tabelle 9: Anzahl der Basenänderungen in einem Triplett, um eine Mutation im T4 Lysozym auszulösen

ID	Mutation	1 Schritt nötig für Mutation	2 Schritte nötig für Mutation	3 Schritte nötig für Mutation
3C7Y	R96A	keine	CGC → GCC	CGC → GCU CGC → GCG CGC → GCA
3C80	R96Y	keine	CGC → UAC	CGC → UAU
3C81	K85A	keine	AAA → GCA	AAA → GCC AAA → GCU AAA → GCG
3C83	D89A	GAU → GCU	GAU → GCC GAU → GCG GAU → GCA	keine
3CDR	R96Q	keine	CGC → CAA CGC → CAG	keine
3C8R	R96G	CGC → GGC	CGC → GGG CGC → GGA CGC → GGU	keine
3C8S	R96E	keine	keine	CGC → GAA CGC → GAG
3CDQ	R96S	CGC → AGC	CGC → AGU CGC → UCC	CGC → UCU CGC → UCA CGC → UCG
3C7W	R96K	keine	keine	CGC → AAA CGC → AAG
3F8V	R96H	CGC → CAC	CGC → CAU	keine

A.2 RMSD-Diagramme

A.2.1 Bacteriorhodopsin

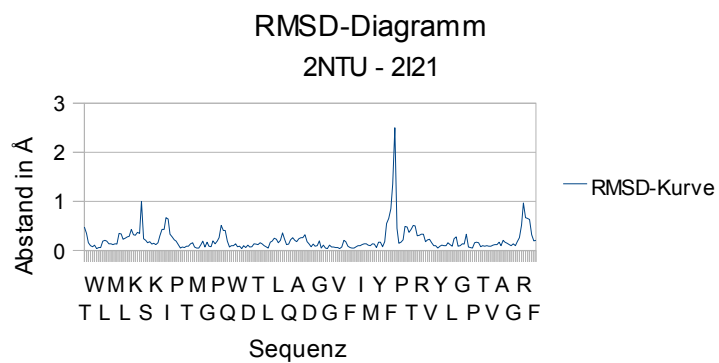


Abbildung 38: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(2I21)

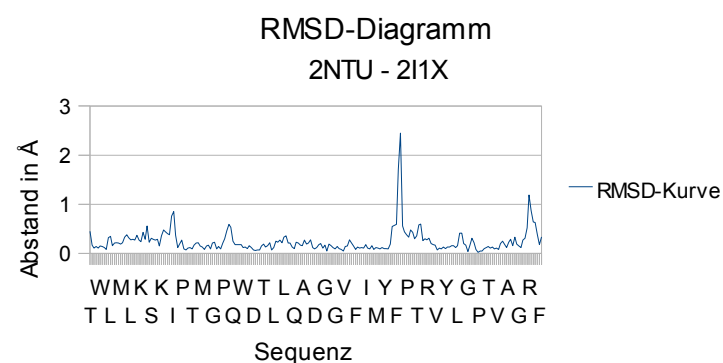


Abbildung 39: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(2I1X)

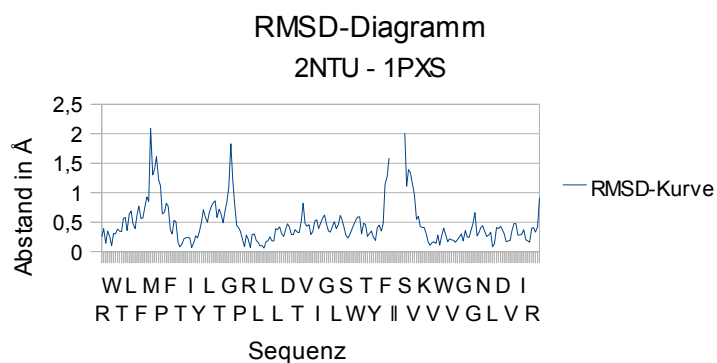


Abbildung 40: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(1PXS)

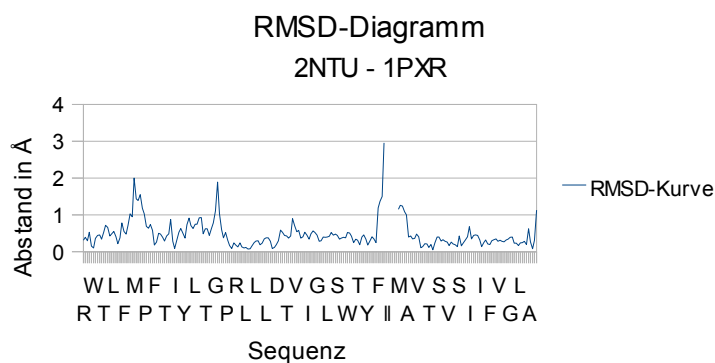


Abbildung 41: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(1PXR)

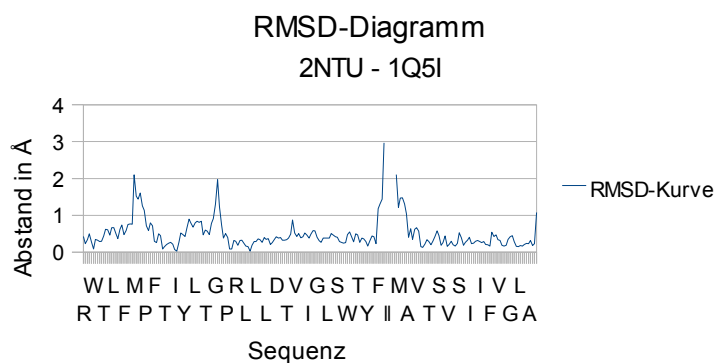


Abbildung 42: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(1Q5I)

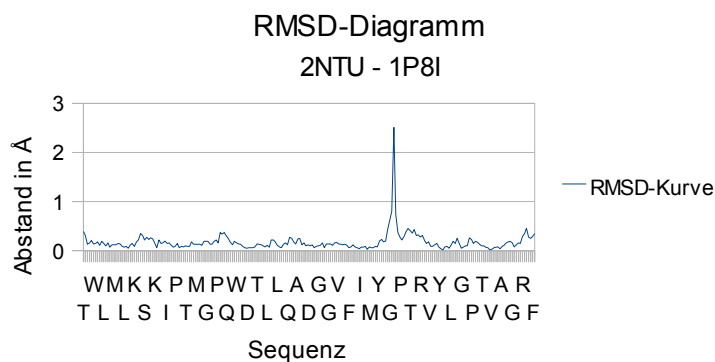


Abbildung 43: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(1P8I)

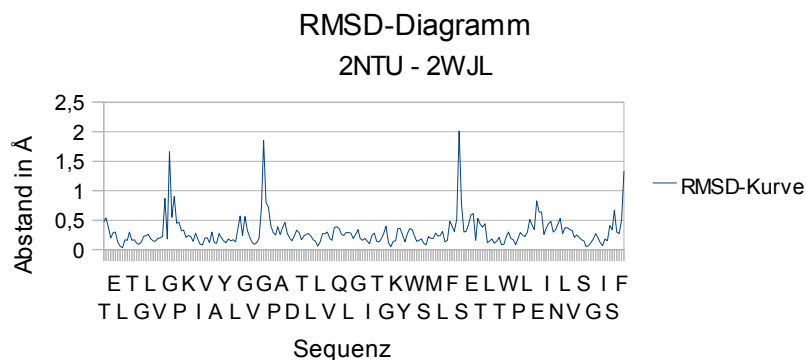


Abbildung 44: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(2WJL)

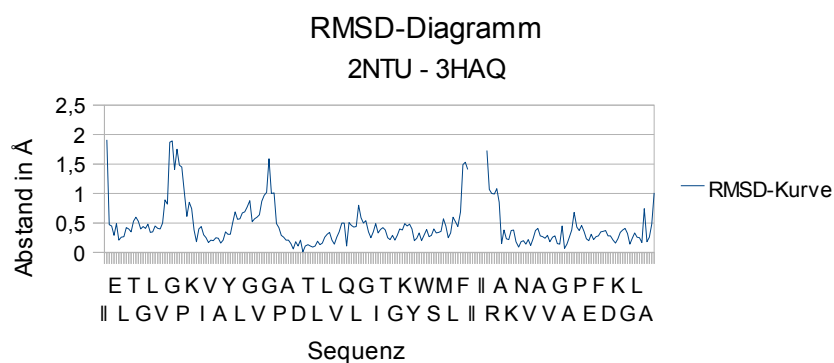


Abbildung 45: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(3HAQ)

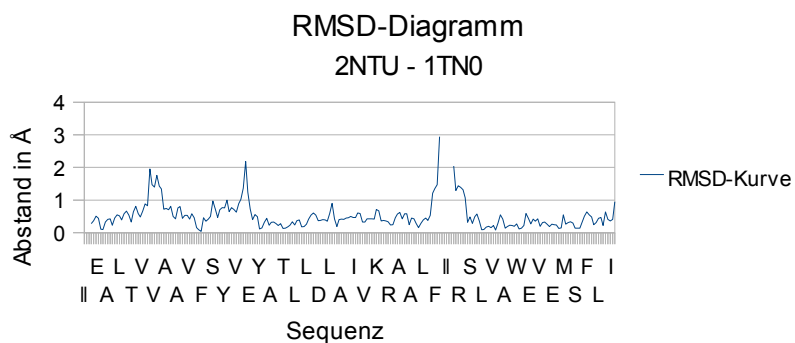


Abbildung 46: RMSD-Diagramm: Wildtyp(2NTU)-Mutante(1TN0)

A.2.2 T4 Lysozym

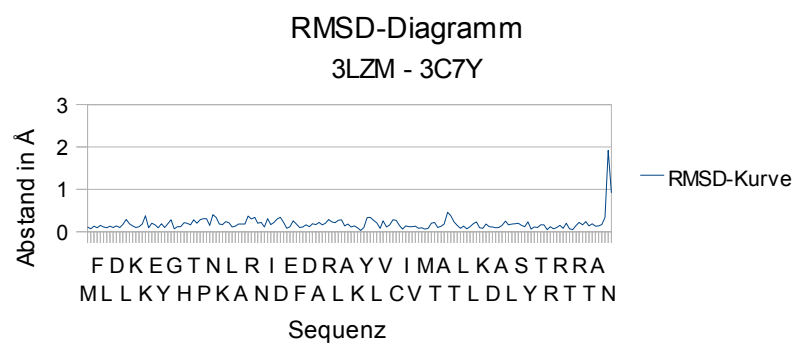


Abbildung 47: RMSD-Diagramm: Wildtyp(3LZM)-Mutante(3C7Y)

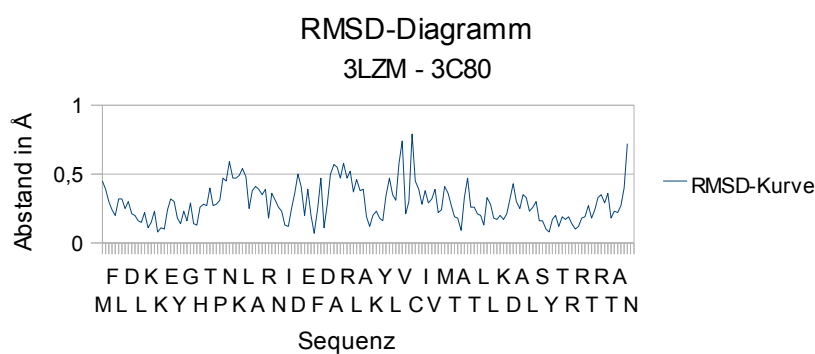


Abbildung 48: RMSD-Diagramm: Wildtyp(3LZM)-Mutante(3C80)

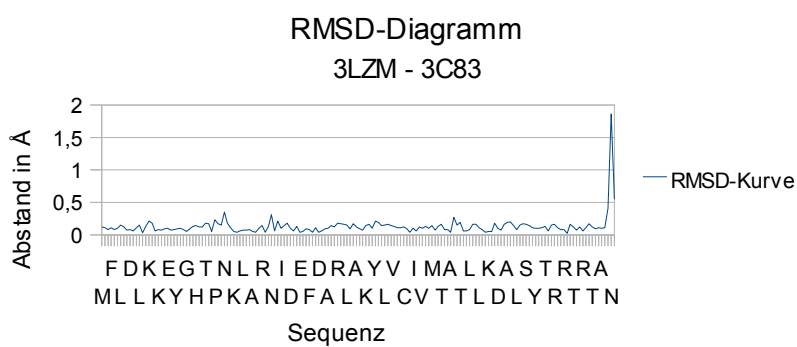


Abbildung 49: RMSD-Diagramm: Wildtyp(3LZM)-Mutante(3C83)

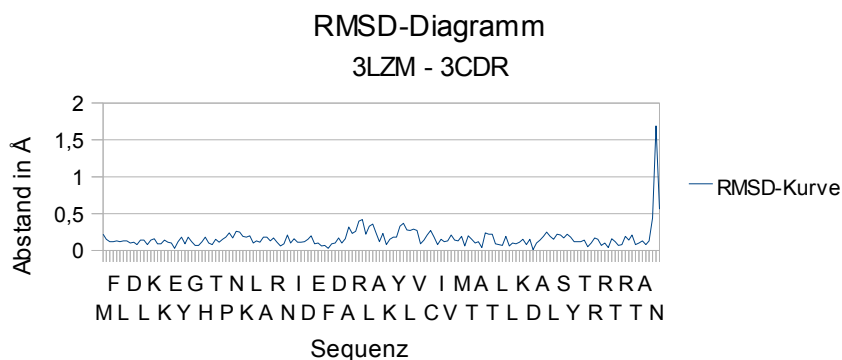


Abbildung 50: RMSD-Diagramm: Wildtyp(3LZM)-Mutante(3CDR)

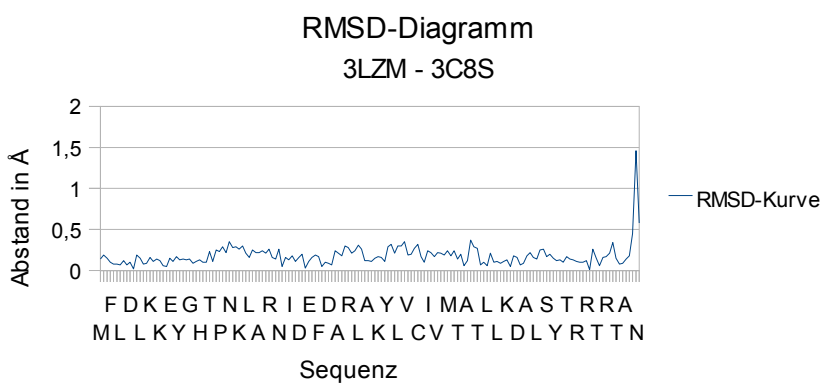


Abbildung 51: RMSD-Diagramm: Wildtyp(3LZM)-Mutante(3C8S)

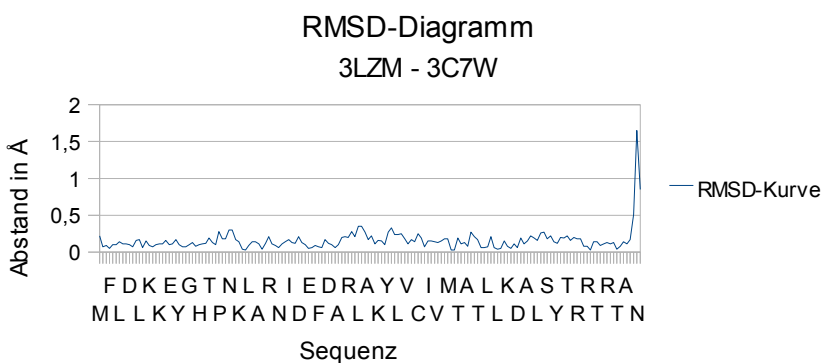


Abbildung 52: RMSD-Diagramm: Wildtyp(3LZM)-Mutante(3C7W)

A.3 Energieprofilalignments

A.3.1 Bacteriorhodopsin

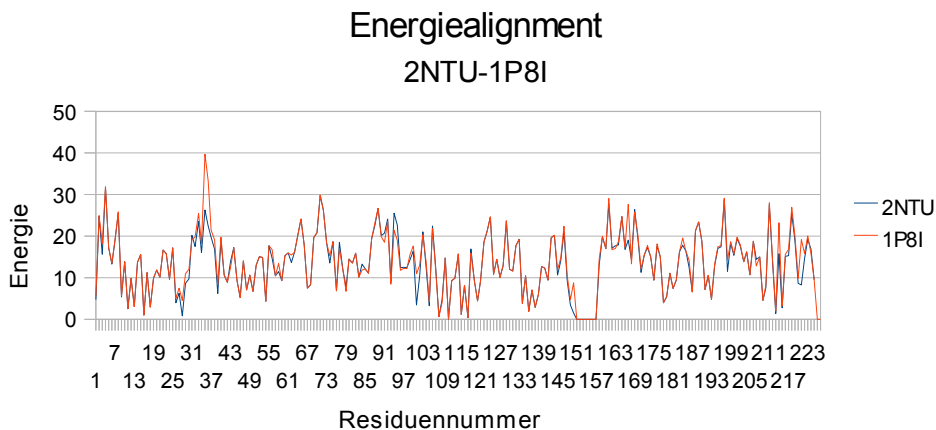


Abbildung 53: Energieprofilalignment: Wildtyp(2NTU)-Mutante(1P8I)

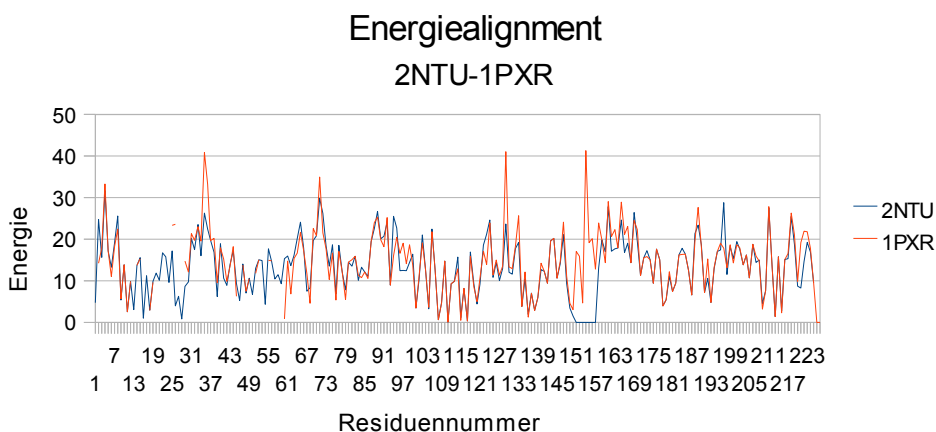


Abbildung 54: Energieprofilalignment: Wildtyp(2NTU)-Mutante(1PXR)

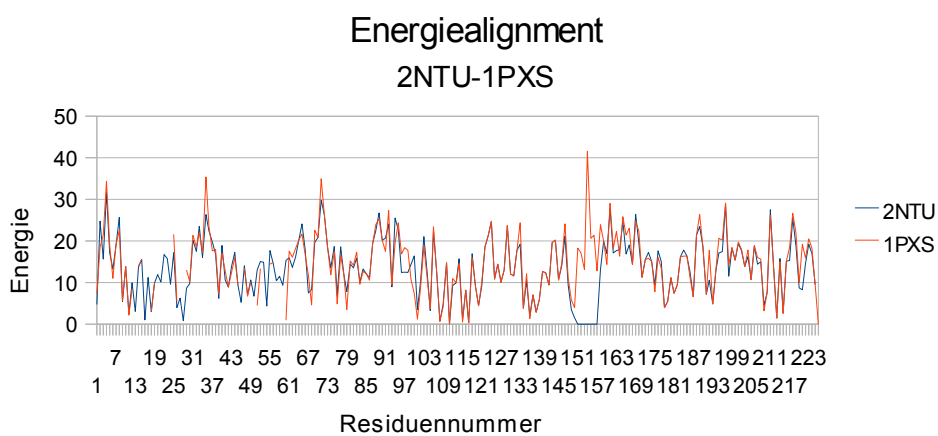


Abbildung 55: Energieprofilalignment: Wildtyp(2NTU)-Mutante(1PXS)

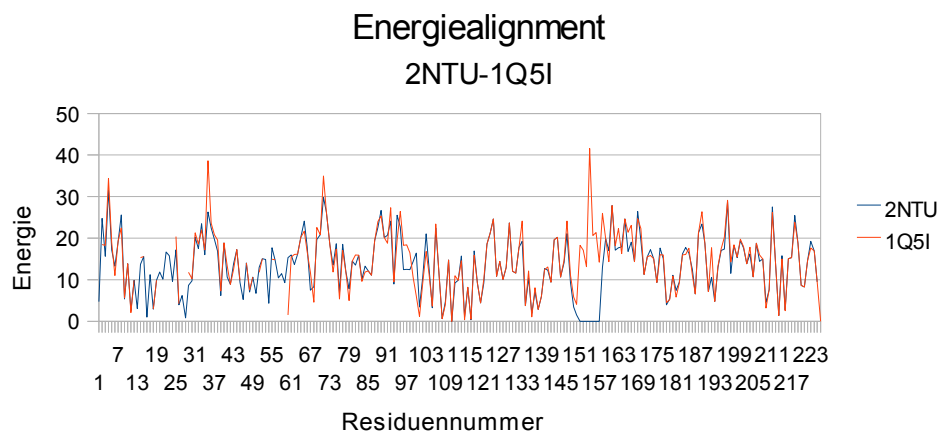


Abbildung 56: Energieprofilalignment: Wildtyp(2NTU)-Mutante(1Q5I)

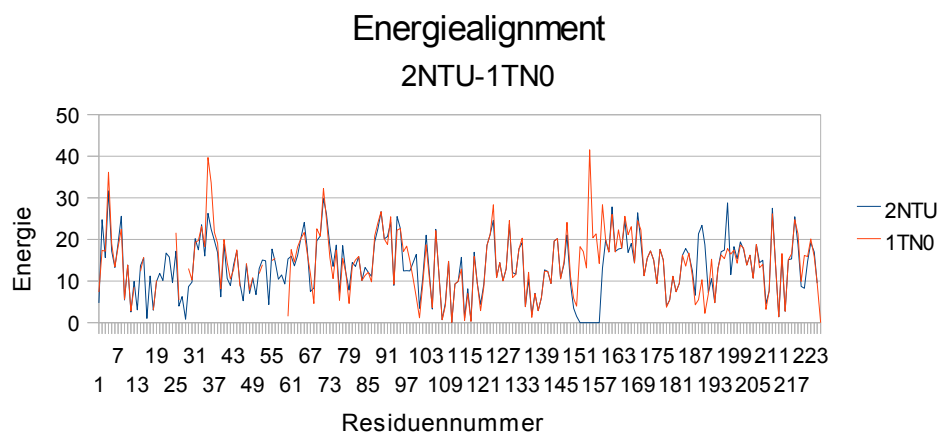


Abbildung 57: Energieprofilalignment: Wildtyp(2NTU)-Mutante(1TN0)

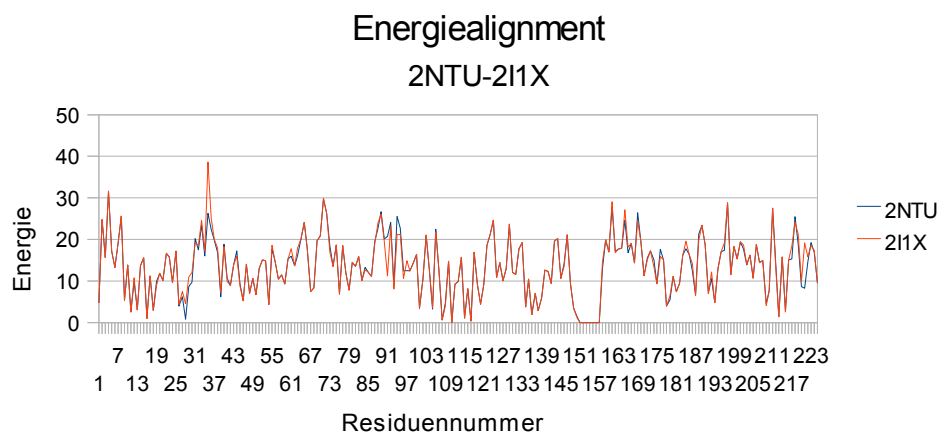


Abbildung 58: Energieprofilalignment: Wildtyp(2NTU)-Mutante(2I1X)

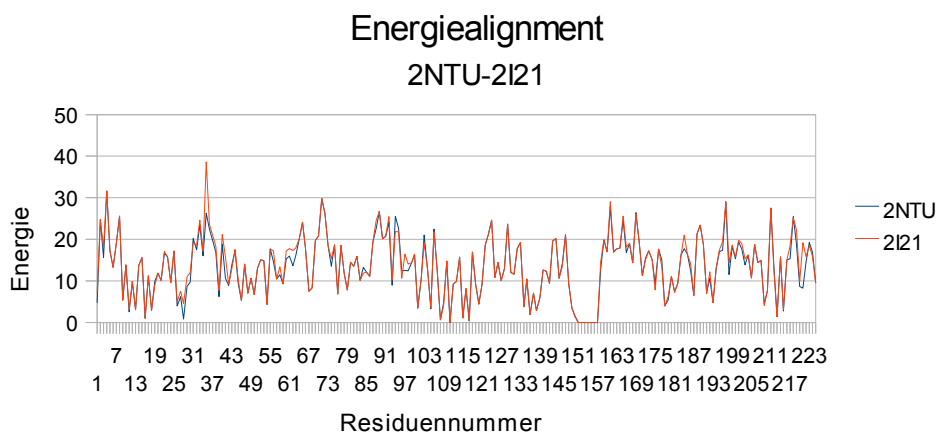


Abbildung 59: Energieprofilalignment: Wildtyp(2NTU)-Mutante(2I21)

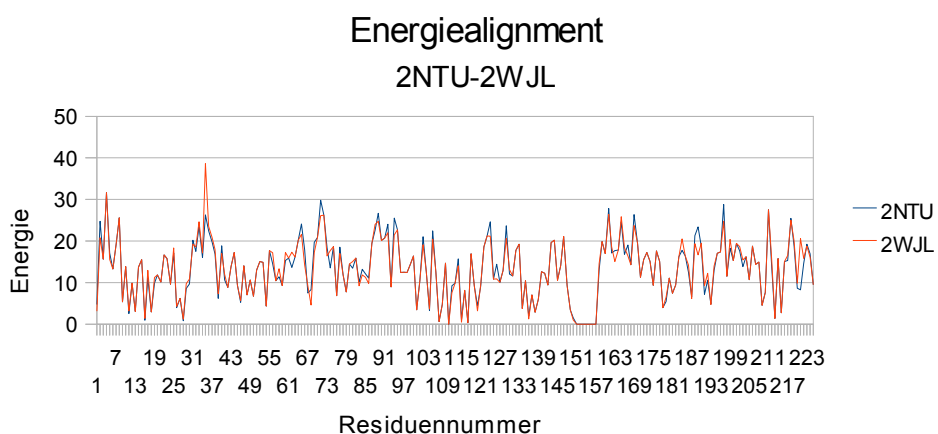


Abbildung 60: Energieprofilalignment: Wildtyp(2NTU)-Mutante(2WJL)

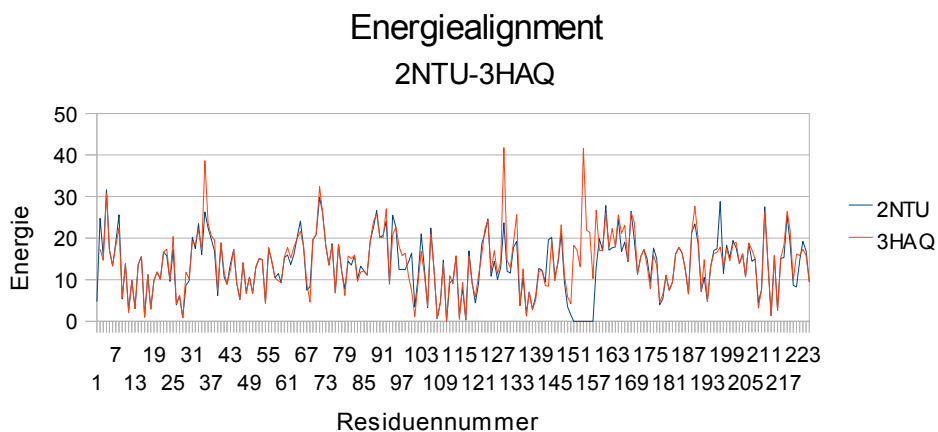


Abbildung 61: Energieprofilalignment: Wildtyp(2NTU)-Mutante(3HAQ)

A.3.2 T4 Lysozym

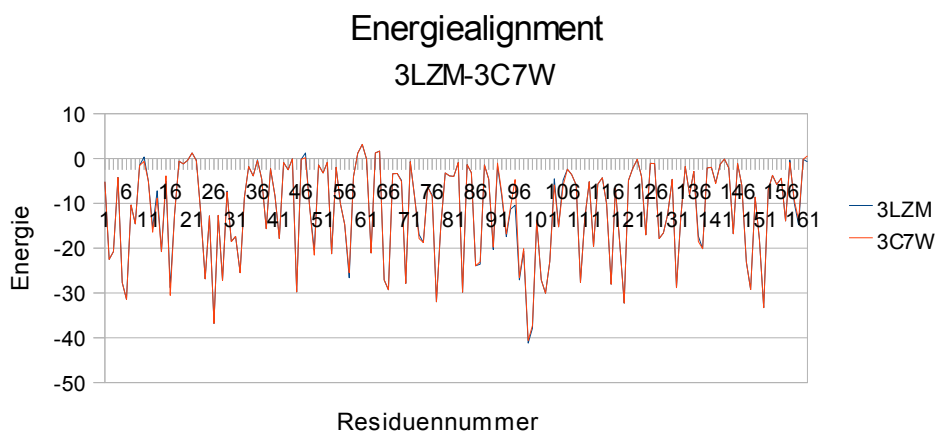


Abbildung 62: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3C7W)

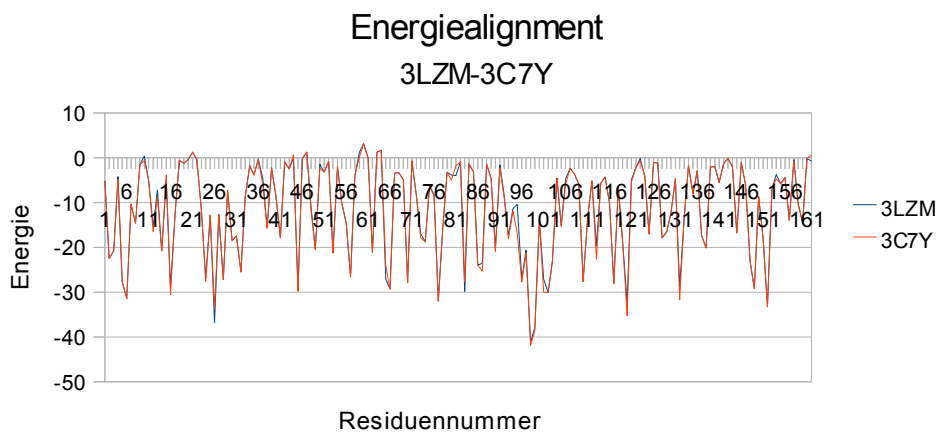


Abbildung 63: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3C7Y)

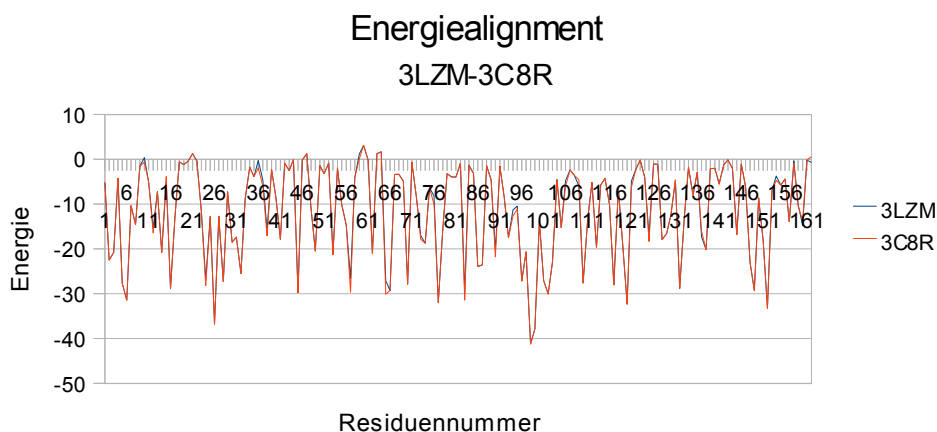


Abbildung 64: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3C8R)

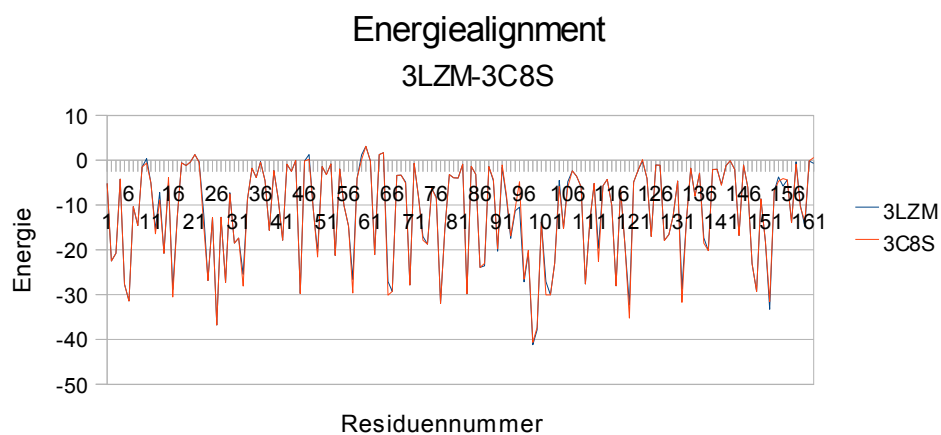


Abbildung 65: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3C8S)

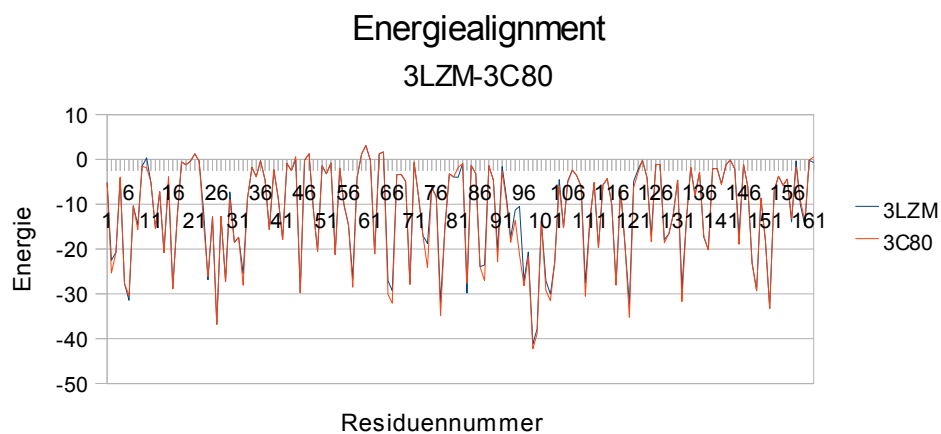


Abbildung 66: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3C80)

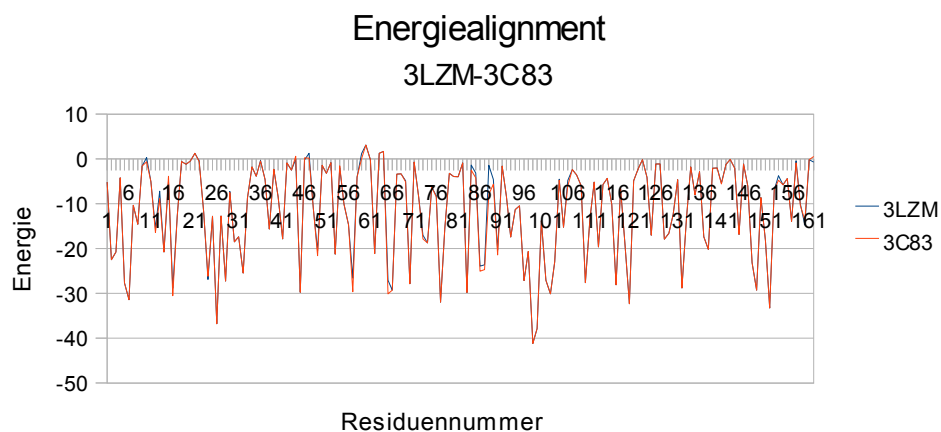


Abbildung 67: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3C83)

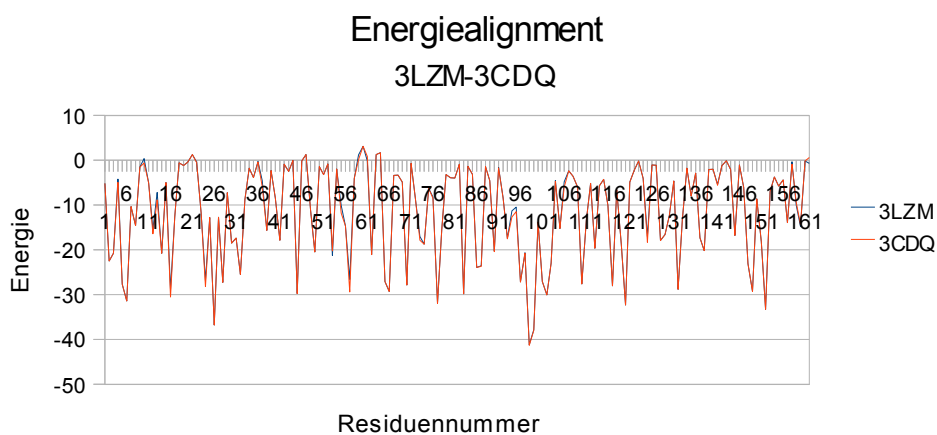


Abbildung 68: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3CDQ)

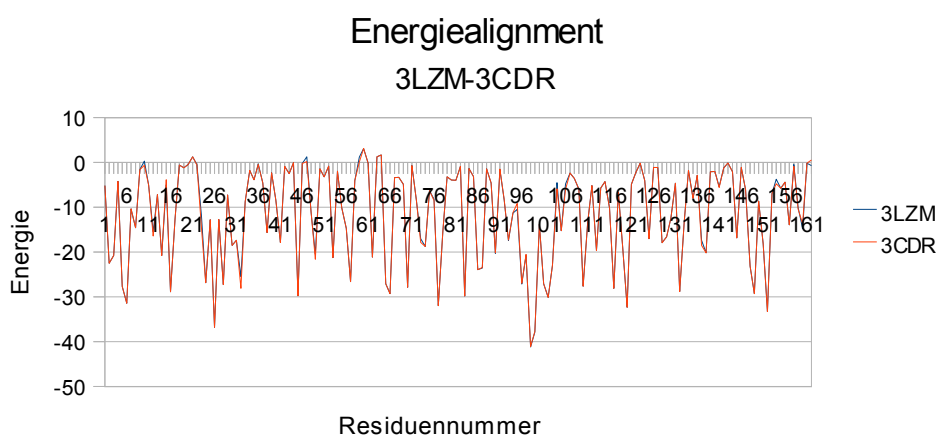


Abbildung 69: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3CDR)

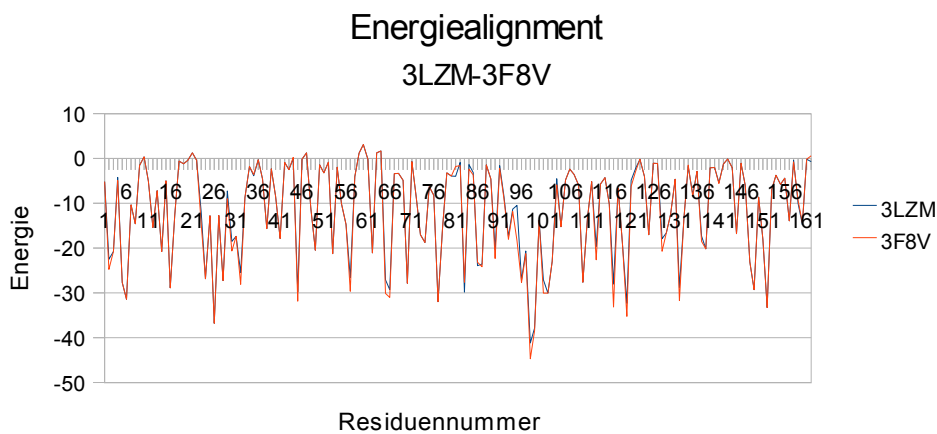


Abbildung 70: Energieprofilalignment: Wildtyp(3LZM)-Mutante(3F8V)

A.4 Mutationsenergieprofil

A.4.1 Bacteriorhodopsin

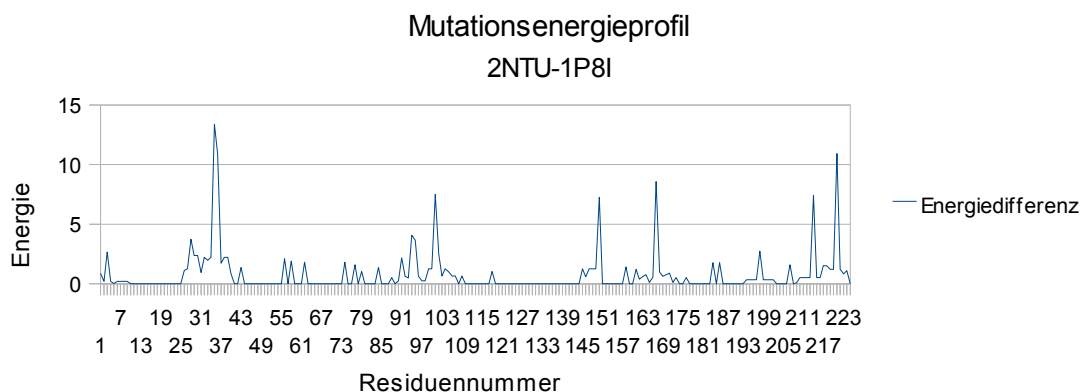


Abbildung 71: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(1P8I)

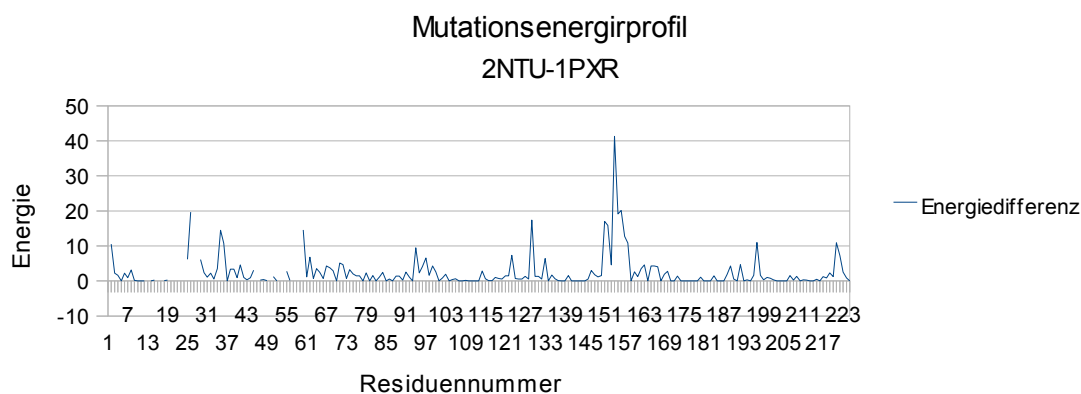


Abbildung 72: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(1PXR)

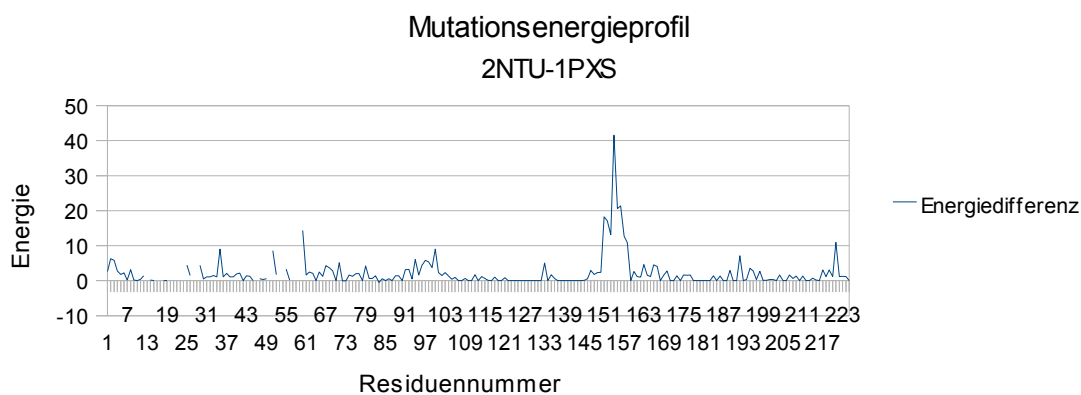


Abbildung 73: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(1PXS)

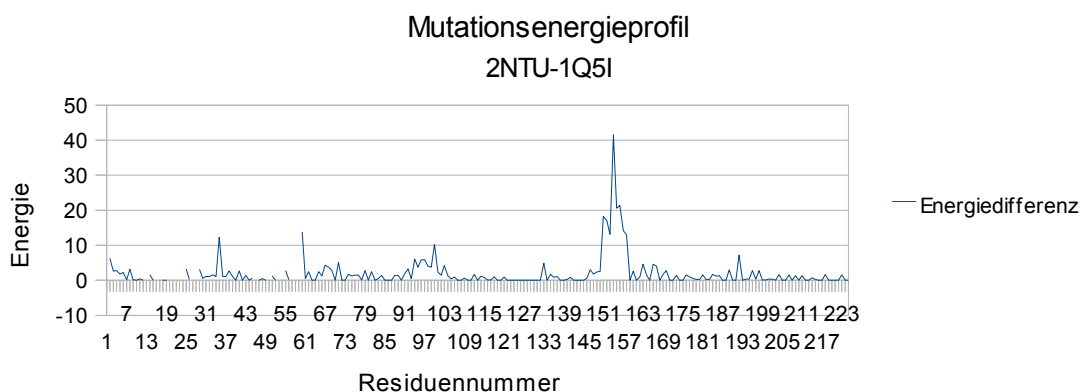


Abbildung 74: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(1Q5I)

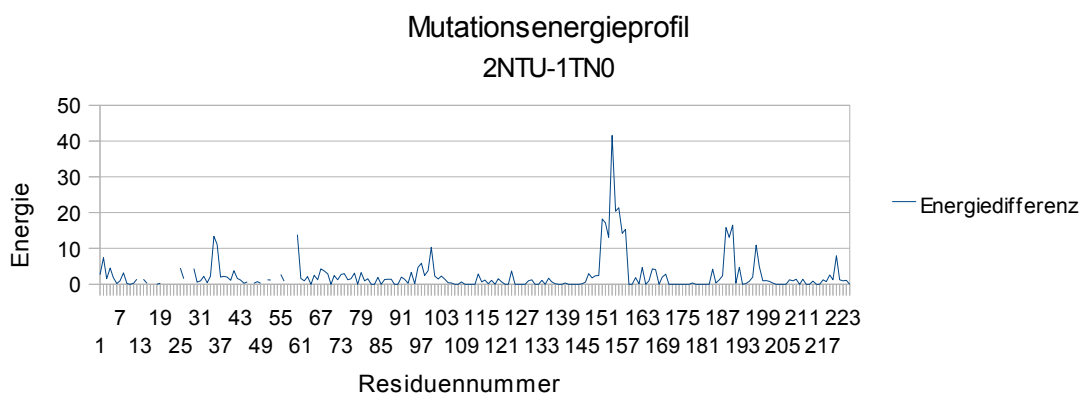


Abbildung 75: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(1TN0)

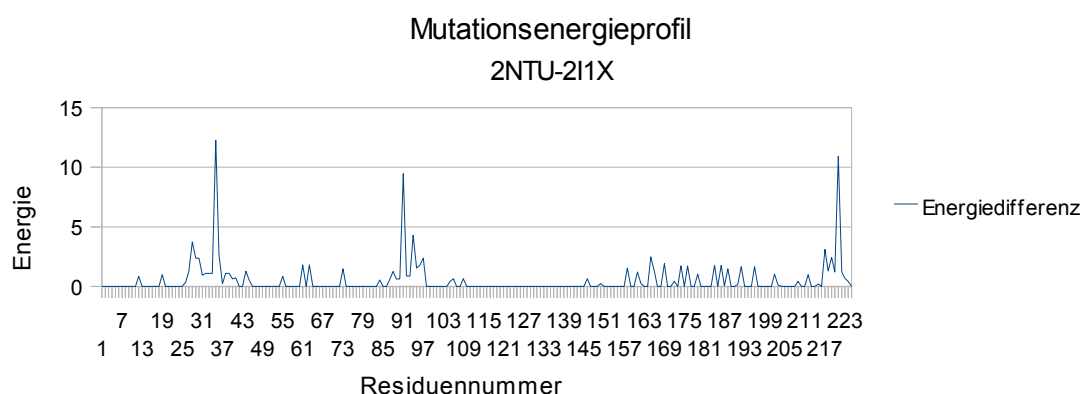


Abbildung 76: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(2I1X)

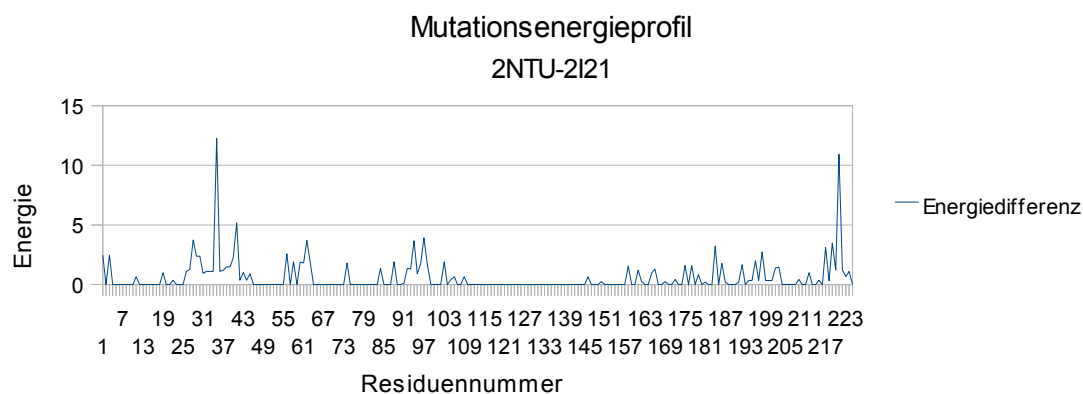


Abbildung 77: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(2I21)

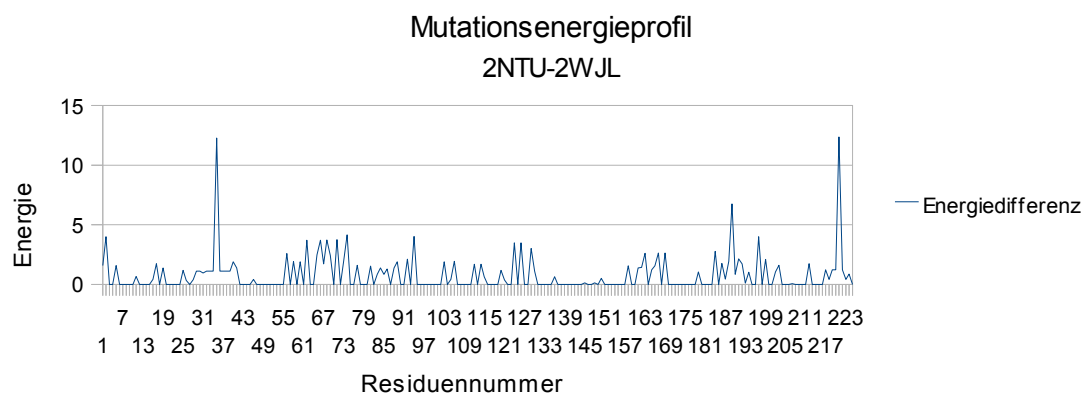


Abbildung 78: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(2WJL)

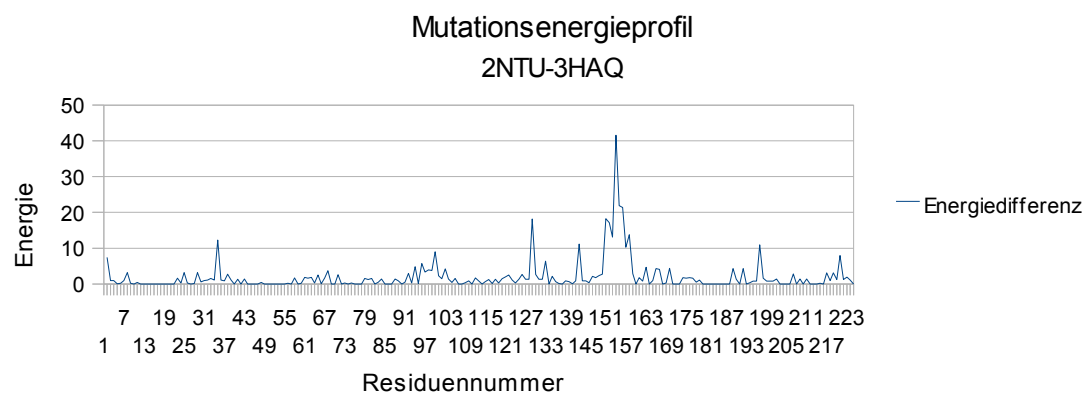
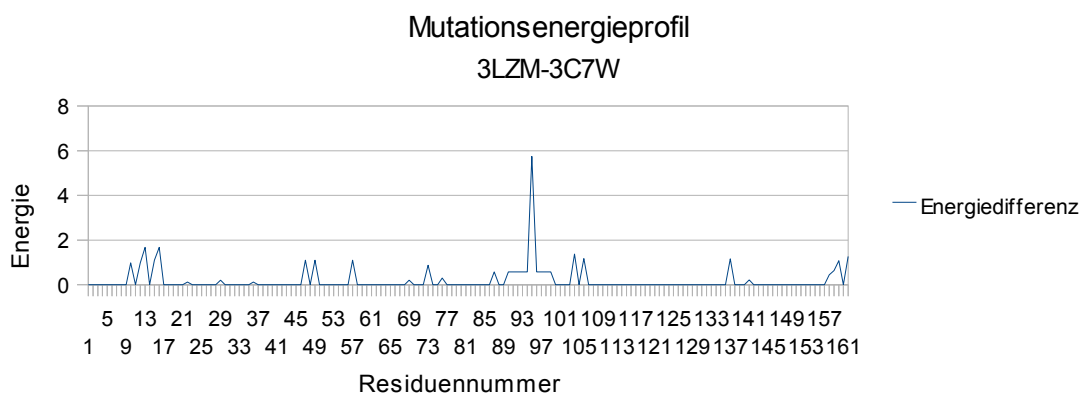
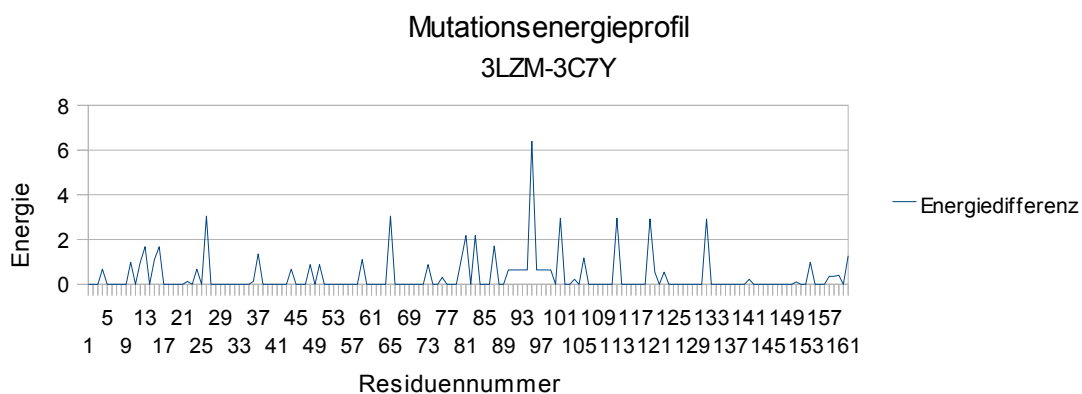
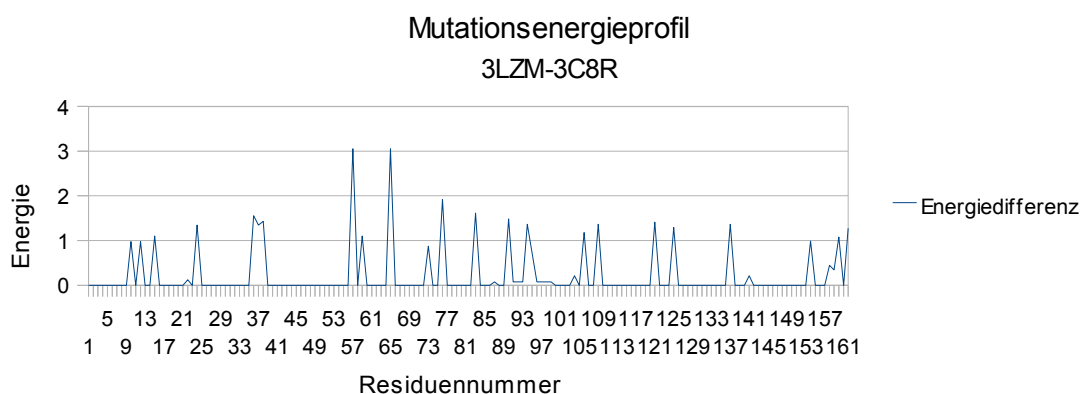


Abbildung 79: Mutationsenergieprofil: Wildtyp(2NTU)-Mutante(3HAQ)

A.4.2 T4 Lysozym*Abbildung 80: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3C7W)**Abbildung 81: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3C7Y)**Abbildung 82: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3C8R)*

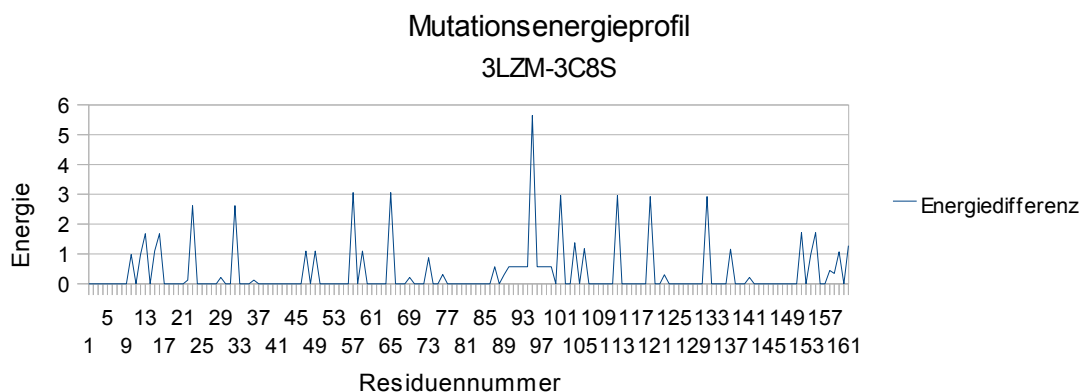


Abbildung 83: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3C8S)

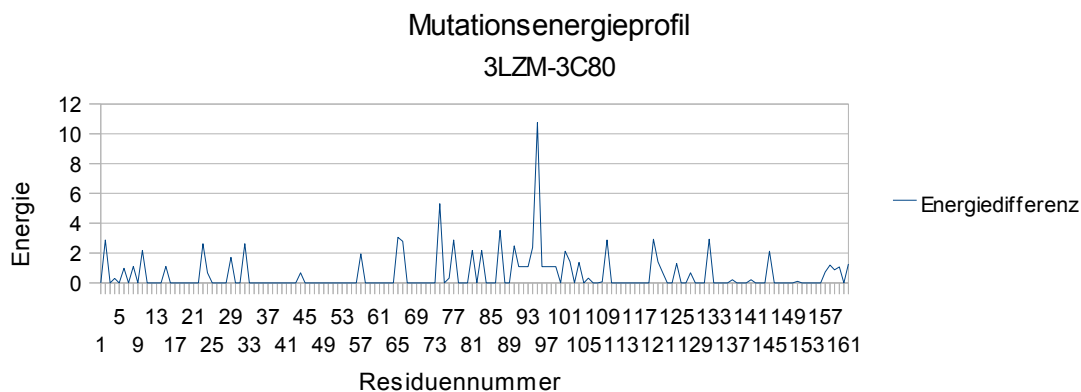


Abbildung 84: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3C80)

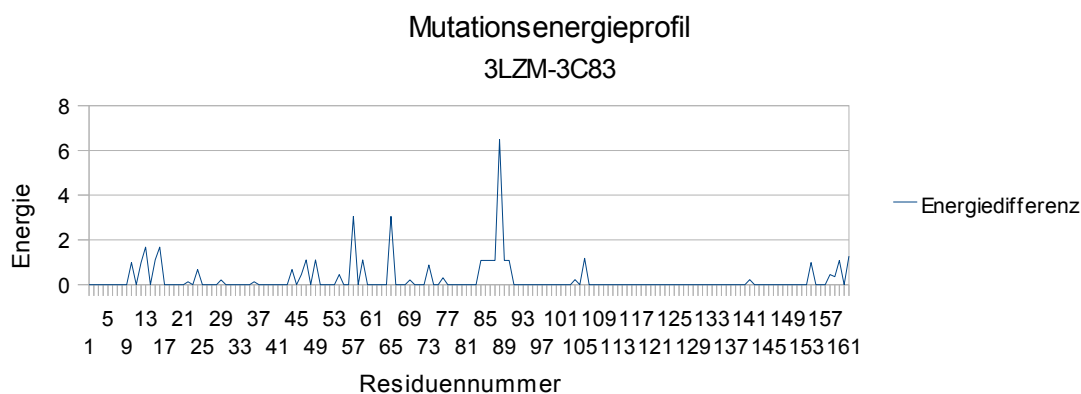


Abbildung 85: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3C83)

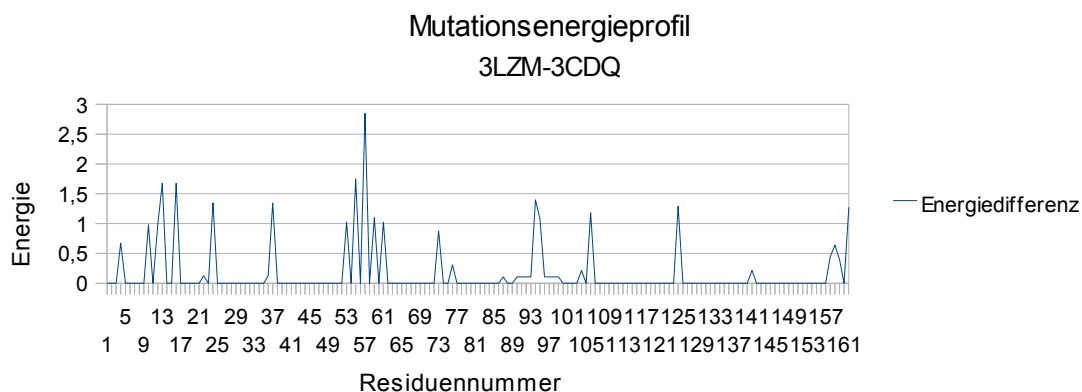


Abbildung 86: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3CDQ)

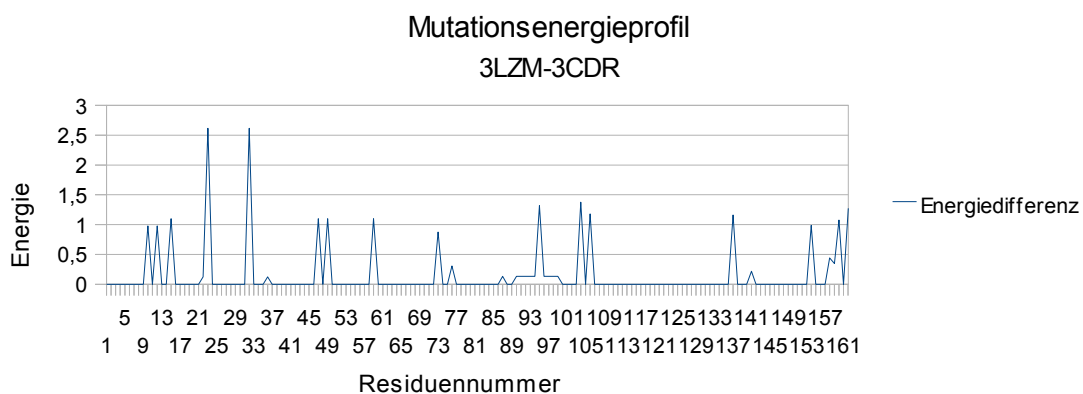


Abbildung 87: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3CDR)

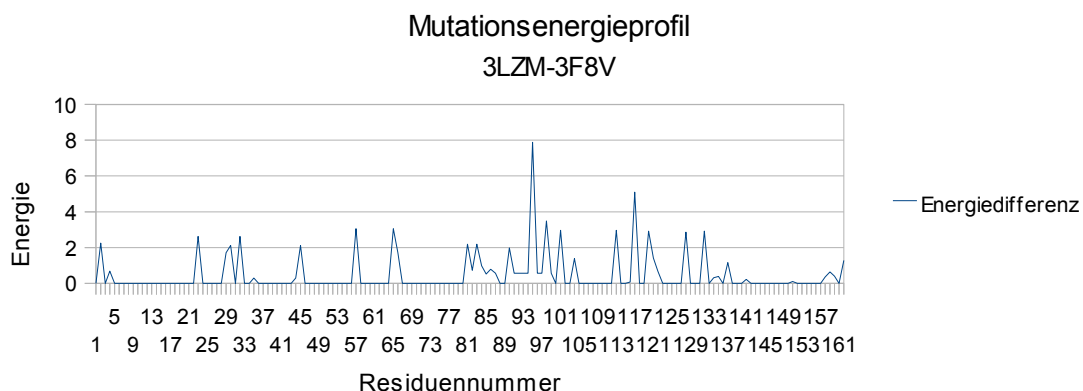


Abbildung 88: Mutationsenergieprofil: Wildtyp(3LZM)-Mutante(3F8V)

Literaturverzeichnis

- [Albrecht, 2011] Albrecht, Urs < urs.albrecht@unifr.ch>: Dreidimensionale Struktur von Proteinen. URL: <http://www.unifr.ch/biochem/assets/files/albrecht/cours/7_2n3_FibrousGlobular.pdf>; verfügbar am 29.03.2011
- [Alexander, 2009] Alexander, P.A., et al.: A minimal sequence code for switching protein structure and function. In: *Proc Natl Acad Sci U S A.* - USA: Center for Advanced Research in Biotechnology. -106(2009)50, S. 21149-54.
- [Behre, 2002] Behre, Jörn Arnold: Analyse von Konformationsänderungen bei der Komplexbildung von Proteinen. -2002.- 106 S. Köln, Universität zu Köln, Mathematisch-Naturwissenschaftliche Fakultät, Diplomarbeit, 2002
- [Biochem, 2011] Unbekannt: Strukturbildende Kräfte in Proteinen. URL: <<http://bc.biochemtech.uni-halle.de/lehre/struct/02-kraefte.pdf>>; verfügbar am 09.08.2011
- [Bolwien, 2002] Bolwien, Carsten: Zeitauflösende Schwingungsspektroskopie an Bacteriorhodopsin und Halorhodopsin. -2002.- 134 S., Düsseldorf, Forschungszentrum Jülich, Biologische Informationsverarbeitung/ Biologische Strukturforschung, Bericht, 2002
- [Brumm, 2010] Brumm, Riccardo: Energieprofilbasierende Stabilitätsanalyse von Membranproteinen. -2010.- 65 S. Mittweida, Hochschule Mittweida, Mathematik/Naturwissenschaften/Informatik, Bachelorarbeit, 2010
- [Böhm, 2005] Böhm, Gerald: T4-Lysozym und Protein Design. URL: <http://www2.biochemtech.uni-halle.de/proteintech/DE/lectures/06_T4_Lysozym.pdf>; verfügbar am 29.03.2011
- [Collins, 2005] Collins, Marcus D.; et al.: Cooperative water filling of a non-polar protein cavity observed by high-pressure crystallography and simulation. In: PNAS. - USA: National Academy of Sciences. - 102(2005)46, S. 16668–16671

- [Cronk, 2011] Cronk, Jeff D. <cronk@gonzaga.edu>: bacteriorhodopsin. URL:<<http://guweb2.gonzaga.edu/faculty/cronk/biochem/B-index.cfm?definition=bacteriorhodopsin>>; verfügbar am 26.04.2011
- [Deplanque, 2011] Deplanque, René <info@fiz-chemie.de>: Membranproteine und ihre Besonderheiten. URL:<http://www.chemgapedia.de/vsengine/vlu/vsc/de/ch/8/bc/vlu/proteine/membranproteine.vlu/Page/vsc/de/ch/8/bc/proteine/funktion_v_prot/membranprot/index.vscml.html>; verfügbar am 28.03.2011
- [Dobson, 1999] Dobson, C. M., A. Šali und M. Karplus: Proteinfaltung aus theoretischer und experimenteller Sicht. Angew Chem, 110(7):908-35, Mar 12 1999
- [Domain, 2011] DomainLoc.com GmbH <publication@DomainLoc.de>: Atombau und Chemische Bindung. URL: <<http://atomaufbau.de/atom05.html>>; verfügbar am 19.07.2011
- [Dressel, 2008] Dressel, Frank: Sequenz, Energie, Struktur – Untersuchungen zur Beziehung zwischen Primär- und Tertiärstruktur in globulären und Membran-Proteinen. -2008.- 101 S. Dresden, Technische Universität, Mathematisch / Naturwissenschaftliche Fakultät, Dissertation, 2008
- [Essen, 1998] Essen, L.-O.; et al.: Lipid Patches in Membrane Proteins: Crystal Structure of the Bacteriorhodopsin-Lipid Complex. Proc. Natl. Acad. Sci. USA 95, 11673-11678 (1998)
- [Faham, 2004] Faham, Salem; et al.: Side-chain Contributions to Membrane Protein Structure and Stability. In: Journal of Molecular Biology. -USA: ELSEVIER. -335(2004)1, S. 297–305
- [Fischer, 2010] Fischer, Lars: Was ist eine Wasserstoffbrückenbindung? In: Angewandte Chemie International Edition. -Heidelberg: Spektrum der Wissenschaft. -1002(2010)10
- [Foobar, 2006] Foobar: Monotopic membrane protein. URL: <http://de.wikibooks.org/wiki/Datei:Monotopic_membrane_protein.png>; verfügbar am 28.03.2011
- [François, 2002] François, Bry <bry@lmu.de>; Peer Kröger: Aktuelles Schlagwort "Datenbanken in der Bioinformatik".

- URL: <<http://www.pms.ifi.lmu.de/publikationen/PMS-FB/PMS-FB-2002-14.pdf>>, verfügbar am 15.06.2011
- [Guerra, 2003] Guerra, Concettina; Istrail, Sorin: Mathematical Methods for Protein Structure Analysis and Design. - 1. Aufl.- Berlin: Springer, 2003
- [Heinke, 2011] Heinke, Florian: Zusammenfassung MEPAL. -2011.- 8 S. Mittweida, Hochschule Mittweida, 2011
- [Helmich, 2011] Helmich, Ulrich <Ulrich.Helmich@t-online.de>: Biomembranen. URL: <<http://www.u-helmich.de/bio/zell/2/23/23.html>>; verfügbar am 21.03.2011
- [Jacobsen, 2005] Jacobsen, Kerstin: Untersuchung der Struktur und Dynamik von T4 Lysozym auf planaren Oberflächen mittels ESR-Spektroskopie. -2005.- 191 S. Berlin, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftlichen Fakultät I, Dissertation, 2005
- [Joh, 2009] Joh, Nathan H.; et al.: Similar Energetic Contributions of Packing in the Core of Membrane and Water Soluble Proteins. In: Journal of the American Chemical Society. - USA: ACS PUBLICATIONS. -131(2009)31, S. 10846–10847
- [Kandt, 2003] Kandt, Christian: Theoretische Untersuchungen integraler photosynthetischer Membranproteine. -2003.- 153 S. Bochum, Ruhr-Universität, Fakultät für Biologie, Dissertation, 2003
- [Koolman, 1998] Koolman, J.; Röhm, K.-H. *Taschenatlas der Biochemie*; Georg Thieme Verlag: Stuttgart, 1998; S. 65
- [Lanyi et al., 2006] Lanyi, Janos K.; Schobert, Brigitte: Propagating Structural Perturbation Inside Bacteriorhodopsin: Crystal Structures of the M State and the D96A and T46V Mutants. In: Biochemistry. - USA: ACS PUBLICATIONS. -45(2006)39, S. 12003–12010
- [Liebbert, 1986] Liebbert, Eike et al.: Allgemeine Biologie. -5.Aufl.- Jena: VEB Gustav Fischer Verlag Jena, 1986

- [Madigan, 2006] M. Madigan, J. Martinko: Brock Mikrobiologie. - 11. Aufl.- München: Pearson Studium, 2006
- [Maurer, 2006] Maurer, Harald: Das Abstoßungsprinzip. URL: <<http://www.mahag.com/grav/abst.php>>; verfügbar am 21.07.2011
- [Merkl, 2009] Merkl, Rainer; Waack, Stephan: Bioinformatik Interaktiv. - 2. Aufl.- Weinheim: WILEY-VCH, 2009
- [Mooers, 2009] Mooers, Blaine H. M.; et al.: Contributions of all 20 amino acids at site 96 to the stability and structure of T4 lysozyme. In: PROTEIN SCIENCE. - Oregon: Wiley-Blackwell. -100(2009)18, S.871—880
- [MooersII, 2009] Mooers, Blaine H. M.; Tronrud, Dale E. and Matthews, Brian W.: Evaluation at atomic resolution of the role of strain in destabilizing the temperature-sensitive T4 lysozyme mutant Arg 96 → His. In: PROTEIN SCIENCE. - Oregon: Wiley-Blackwell. -100(2009)18, S.863—870
- [Oesterhelt, 1971] Oesterhelt, D. und W. Stoeckenius: Rhodopsin-Litre Protein from the Purple Membrane of Halobacterium Halobium. Nat. New Biol., 233(39) :149-152, 1971
- [Oesterhelt I, 1971] Oesterhelt, D. : The Purple Membrane of Halobacterium Halobium: A New System for Light Energy Conversion. Ciba Found Symp, 31:147-167, 1975
- [Pews, 2010] Pews-Hocke, Christa; et al.: Basiswissen Biologie. - 4. Aufl. - Mannheim: Bibliographisches Institut, 2010
- [Rau et al., 2010] Rau et al.: Proteindatenbanken und Strukturanalyse von Proteinen am Computer. URL: <http://www.biophys.uni-frankfurt.de/~wille/prakt/anleitungen/09_proteindatenbanken.pdf>; verfügbar am 29.06.2011
- [Schmieder, 2004] Schmieder, Robert; Janowski, Sebastian: Proteindatenbanken Swiss-Prot & PDB. URL: <http://www.freewebs.com/ician/Proteindatenbanken_Ausarbeitung.pdf>; verfügbar am 29.06.2011

- [Schneider, 2009] Schneider, Sarah: Phylogenetische Analyse von Stabilisierungsmustern in Membranproteinen auf der Basis von Energieprofilen. -2009.- 56 S. Dresden, Technische Universität, Molekulare Biotechnologie, Bachelorarbeit, 2009
- [Schobert, 2003] Schobert, Brigitte; Brown, Leonid S. and Lanyi, Janos K.: Crystallographic Structures of the M and N Intermediates of Bacteriorhodopsin: Assembly of a Hydrogen-bonded Chain of Water Molecules Between Asp-96 and the Retinal Schiff Base. In: Journal of Molecular Biology. -USA: ELSEVIER. -330(2003)1, S. 553–570
- [Voet, 2002] Voet, D.; Voet, J.G.; Pratt, C.W. *Fundamentals of Biochemistry*; John Wiley & Sons, Inc.: New York, 2002; Rev. ed., Chapter 6, 126
- [Wachtveitl, 2011] Wachtveitl, Josef <wveitl@theochem.uni-frankfurt.de>: Der Proteinfaltung auf der Spur. URL: <<http://www.biophys.uni-frankfurt.de/layout/pdf/proteinfaltung.pdf>>; verfügbar am 18.07.2011
- [Yohannan, 2004] Yohannan, Sarah; et al.: Proline Substitutions are not Easily Accommodated in a Membrane Protein. In: Journal of Molecular Biology. - USA: ELSEVIER. -341(2004)1, S. 1-6
- [YohannanII, 2004] Yohannan, Sarah; et al.: The evolution of transmembrane helix kinks and the structural diversity of G protein-coupled receptors. In: Proceedings of the National Academy of Sciences. -USA: National Academy of Sciences. -101(2004)4, S. 959–963

Erklärung zur selbständigen Anfertigung der Arbeit

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur angefertigt habe.

Mittweida, den 22.08.2011

Anja Grajewski